

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2005-102192

(P2005-102192A)

(43) 公開日 平成17年4月14日(2005.4.14)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
H04N 7/24	H04N 7/13	5C053
H04N 5/92	H04N 5/92	5C059

審査請求 未請求 請求項の数 8 O L (全 31 頁)

(21) 出願番号	特願2004-256203 (P2004-256203)	(71) 出願人	000002185 ソニー株式会社 東京都品川区北品川6丁目7番35号
(22) 出願日	平成16年9月2日(2004.9.2)	(74) 代理人	100082740 弁理士 田辺 惠基
(31) 優先権主張番号	特願2003-310639 (P2003-310639)	(72) 発明者	塚越 郁夫 東京都品川区北品川6丁目7番35号ソニー株式会社内
(32) 優先日	平成15年9月2日(2003.9.2)	(72) 発明者	高田 信司 東京都品川区北品川6丁目7番35号ソニー株式会社内
(33) 優先権主張国	日本国(JP)	(72) 発明者	後藤 晃一 東京都品川区北品川6丁目7番35号ソニー株式会社内

最終頁に続く

(54) 【発明の名称】 コンテンツ受信装置、ビデオオーディオ出力タイミング制御方法及びコンテンツ提供システム

(57) 【要約】

【課題】

本発明は、視聴者に違和感を感じさせることなく映像及び音声間のリップシンクをデコーダ側で確実に調整できるようにする。

【解決手段】

本発明は、ビデオタイムスタンプVTSが付された複数の符号化ビデオフレームと、オーディオタイムスタンプATSが付された複数の符号化オーディオフレームとを受信して復号した結果得られる複数のビデオフレームVF1及び複数のオーディオフレームAF1を蓄積し、エンコーダ側の基準クロックとデコーダ側のシステムタイムクロックstcとのずれによって生じる時間差をレンダラー37、67で算出し、その時間差に応じ、複数のオーディオフレームAF1をフレーム単位で順次出力するときのオーディオフレーム出力タイミングを基準として複数のビデオフレームVF1のビデオフレーム出力タイミングをフレーム単位で調整することにより、音声の連続性を保持したままリップシンクさせることができる。

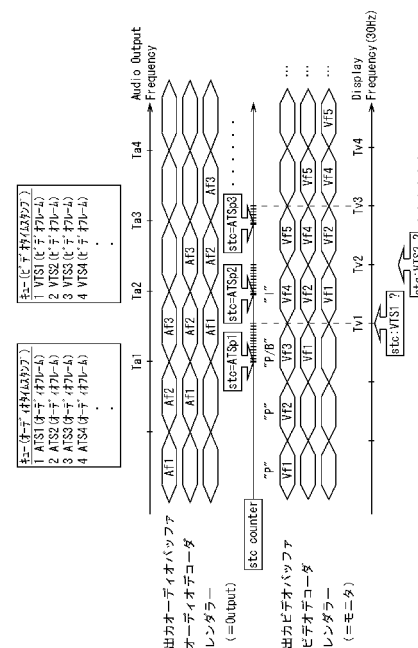


図7 フリエンコーデッドストリーミングにおけるビデオフレーム及びオーディオフレームの出力タイミング

【特許請求の範囲】

【請求項 1】

エンコーダ側の基準クロックに基づくビデオタイムスタンプが順次付された複数の符号化ビデオフレームと、上記基準クロックに基づくオーディオタイムスタンプが順次付された複数の符号化オーディオフレームとを上記エンコーダ側のコンテンツ提供装置から受信して復号する復号手段と、

上記復号手段によって上記符号化ビデオフレーム及び上記符号化オーディオフレームを復号した結果得られる複数のビデオフレーム及び複数のオーディオフレームを蓄積する記憶手段と、

上記エンコーダ側の基準クロックのクロック周波数とデコーダ側のシステムタイムクロックのクロック周波数とのずれによって生じる時間差を算出する算出手段と、

上記時間差に応じ、上記複数のオーディオフレームをフレーム単位で順次出力するときのオーディオフレーム出力タイミングを基準として上記複数のビデオフレームをフレーム単位で順次出力するときのビデオフレーム出力タイミングを調整するタイミング調整手段と

を具えることを特徴とするコンテンツ受信装置。

【請求項 2】

上記タイミング調整手段は、上記時間差が所定の時間よりも短い場合、上記デコーダ側のシステムタイムクロックの基で上記ビデオタイムスタンプに従い上記ビデオフレームを出力する

ことを特徴とする請求項 1 に記載のコンテンツ受信装置。

【請求項 3】

上記コンテンツ提供装置からリアルタイム性が要求される UDP (User Datagram Protocol) で送信される上記エンコーダ側の基準クロックを受信する受信手段

を具え、

上記算出手段は、上記エンコーダ側の基準クロックと上記デコーダ側のシステムタイムクロックとを同期させた上で、上記エンコーダ側の基準クロックのクロック周波数とデコーダ側のシステムタイムクロックのクロック周波数とのずれによって生じる時間差を算出する

ことを特徴とする請求項 1 に記載のコンテンツ受信装置。

【請求項 4】

上記コンテンツ受信装置は、

上記ビデオタイムスタンプ、上記オーディオタイムスタンプの順番で決められたプリセットシーケンスに従い、当該ビデオタイムスタンプ又は当該オーディオタイムスタンプを用いて上記デコーダ側のシステムタイムクロックをプリセットするプリセット手段を具え、

上記算出手段は、上記エンコーダ側の基準クロックのクロック周波数と上記プリセット後のシステムタイムクロックのクロック周波数とのずれによって生じる上記時間差を算出する

ことを特徴とする請求項 1 に記載のコンテンツ受信装置。

【請求項 5】

上記コンテンツ受信装置は、

上記ビデオタイムスタンプ、上記オーディオタイムスタンプの順番で決められたプリセットシーケンスに従うことにより、上記コンテンツの種類がオーディオのみでなるものの場合、上記オーディオタイムスタンプを用いて上記デコーダ側のシステムタイムクロックをプリセットするプリセット手段と、

上記プリセット後のシステムタイムクロック及び上記オーディオタイムスタンプに基づいて上記オーディオフレームの音声を出力するオーディオ出力手段と

を具えることを特徴とする請求項 1 に記載のコンテンツ受信装置。

【請求項 6】

10

20

30

40

50

上記コンテンツ受信装置は、

上記ビデオタイムスタンプ、上記オーディオタイムスタンプの順番で決められたプリセットシーケンスに従うことにより、上記コンテンツの種類がビデオのみでなるものの場合、上記ビデオタイムスタンプを用いて上記デコーダ側のシステムタイムクロックをプリセットするプリセット手段と、

上記プリセット後のシステムタイムクロック及び上記ビデオタイムスタンプに基づいて上記ビデオフレームの映像を出力するビデオ出力手段と

を具えることを特徴とする請求項 1 に記載のコンテンツ受信装置。

【請求項 7】

復号手段に対して、エンコーダ側の基準クロックに基づくビデオタイムスタンプが順次付された複数の符号化ビデオフレームと、上記基準クロックに基づくオーディオタイムスタンプが順次付された複数の符号化オーディオフレームとを上記エンコーダ側のコンテンツ提供装置から受信して復号させる復号ステップと、

記憶手段に対して、上記復号ステップで上記符号化ビデオフレーム及び上記符号化オーディオフレームを復号した結果得られる複数のビデオフレーム及び複数のオーディオフレームを蓄積させる記憶ステップと、

算出手段に対して、上記エンコーダ側の基準クロックのクロック周波数とデコーダ側のシステムタイムクロックのクロック周波数とのずれによって生じる時間差を算出させる差分算出ステップと、

タイミング調整手段に対して、上記時間差に応じ、上記複数のオーディオフレームをフレーム単位で順次出力するときのオーディオフレーム出力タイミングを基準として上記複数のビデオフレームをフレーム単位で順次出力するときのビデオフレーム出力タイミングを調整させるタイミング調整ステップと

を具えることを特徴とするビデオオーディオ出力タイミング制御方法。

【請求項 8】

コンテンツ提供装置とコンテンツ受信装置を有するコンテンツ提供システムであって、上記コンテンツ提供装置は、

エンコーダ側の基準クロックに基づくビデオタイムスタンプを付した複数の符号化ビデオフレームと、上記基準クロックに基づくオーディオタイムスタンプを付した複数の符号化オーディオフレームとを生成する符号化手段と、

上記複数の符号化ビデオフレーム及び上記複数の符号化オーディオフレームを上記コンテンツ受信装置へ順次送信する送信手段と

を具え、

上記コンテンツ受信装置は、

上記ビデオタイムスタンプが順次付された複数の符号化ビデオフレームと、上記オーディオタイムスタンプが順次付された複数の符号化オーディオフレームとを上記エンコーダ側のコンテンツ提供装置から受信して復号する復号手段と、

上記復号手段によって上記符号化ビデオフレーム及び上記符号化オーディオフレームを復号した結果得られる複数のビデオフレーム及び複数のオーディオフレームを蓄積する記憶手段と、

上記エンコーダ側の基準クロックのクロック周波数とデコーダ側のシステムタイムクロックのクロック周波数とのずれによって生じる時間差を算出する算出手段と、

上記時間差に応じ、上記複数のオーディオフレームをフレーム単位で順次出力するときのオーディオフレーム出力タイミングを基準として上記複数のビデオフレームをフレーム単位で順次出力するときのビデオフレーム出力タイミングを調整するタイミング調整手段と

を具えることを特徴とするコンテンツ提供システム。

【発明の詳細な説明】

【技術分野】

【0001】

10

20

30

40

50

本発明は、コンテンツ受信装置、ビデオオーディオ出力タイミング制御方法及びコンテンツ提供システムに関し、例えばコンテンツを受信するデコーダ側で映像と音声のリップシンクがずれることを解消する場合に適用して好適なものである。

【背景技術】

【0002】

従来、コンテンツ受信装置においては、エンコーダ側のサーバからコンテンツを受信してデコードする場合、当該コンテンツを構成する映像パケット及び音声パケットに分離し、それぞれデコードした後に映像パケットに付されたビデオタイムスタンプと音声パケットに付されたオーディオタイムスタンプを基にビデオフレームとオーディオフレームを出力することにより、映像と音声との出力タイミングを一致させる（すなわちリップシンクさせる）ようになされている（例えば、特許文献1、特許文献2参照）。 10

【特許文献1】特開平8-280008号公報

【特許文献2】特開2004-15553公報

【発明の開示】

【発明が解決しようとする課題】

【0003】

ところでかかる構成のコンテンツ受信装置においては、当該デコーダ側のシステムタイムクロックと、エンコーダ側の基準クロックとが互いに同期しているとは限らず、また当該デコーダ側のシステムタイムクロックにおけるクロックジッタ等によってエンコーダ側の基準クロックとの間でクロック周波数の微妙なずれが生じていることもある。 20

【0004】

またコンテンツ受信装置は、ビデオフレームとオーディオフレームとではそのデータ長が異なるため、当該デコーダ側のシステムタイムクロックとエンコーダ側の基準クロックとが完全に同期していないときには、ビデオタイムスタンプ及びオーディオタイムスタンプを基にビデオフレーム及びオーディオフレームを出力したとしても、映像と音声との出力タイミングが一致せず、リップシンクがずれてしまうという問題があった。

【0005】

本発明は以上の点を考慮してなされたもので、視聴者であるユーザに違和感を感じさせることなく映像及び音声間のリップシンクを当該デコーダ側で確実に調整し得るコンテンツ受信装置、ビデオオーディオ出力タイミング制御方法及びコンテンツ提供システムを提案しようとするものである。 30

【課題を解決するための手段】

【0006】

かかる課題を解決するため本発明においては、エンコーダ側の基準クロックに基づくビデオタイムスタンプが順次付された複数の符号化ビデオフレームと、基準クロックに基づくオーディオタイムスタンプが順次付された複数の符号化オーディオフレームとをエンコーダ側のコンテンツ提供装置から受信して復号する復号手段と、復号手段によって符号化ビデオフレーム及び符号化オーディオフレームを復号した結果得られる複数のビデオフレーム及び複数のオーディオフレームを蓄積する記憶手段と、エンコーダ側の基準クロックのクロック周波数とデコーダ側のシステムタイムクロックのクロック周波数とのずれによって生じる時間差を算出する算出手段と、時間差に応じ、複数のオーディオフレームをフレーム単位で順次出力するときのオーディオフレーム出力タイミングを基準として複数のビデオフレームをフレーム単位で順次出力するときのビデオフレーム出力タイミングを調整するタイミング調整手段とを設けるようにする。 40

【0007】

エンコーダ側の基準クロックとデコーダ側のシステムタイムクロックとの間におけるクロック周波数のずれによって生じる時間差に応じ、複数のオーディオフレームをフレーム単位で順次出力するときのオーディオフレーム出力タイミングを基準として複数のビデオフレームをフレーム単位で順次出力するときのビデオフレーム出力タイミングを調整することにより、エンコーダ側とデコーダ側のクロック周波数の差を吸収し、オーディオフレ 50

ーム出力タイミングにビデオフレーム出力タイミングを合わせてリップシンクさせることができる。

【0008】

また本発明においては、復号手段に対して、エンコーダ側の基準クロックに基づくビデオタイムスタンプが順次付された複数の符号化ビデオフレームと、基準クロックに基づくオーディオタイムスタンプが順次付された複数の符号化オーディオフレームとをエンコーダ側のコンテンツ提供装置から受信して復号させる復号ステップと、記憶手段に対して、復号ステップで符号化ビデオフレーム及び符号化オーディオフレームを復号した結果得られる複数のビデオフレーム及び複数のオーディオフレームを蓄積させる記憶ステップと、算出手段に対して、エンコーダ側の基準クロックのクロック周波数とデコーダ側のシステムタイムクロックのクロック周波数とのずれによって生じる時間差を算出させる差分算出ステップと、タイミング調整手段に対して、時間差に応じ、複数のオーディオフレームをフレーム単位で順次出力するときのオーディオフレーム出力タイミングを基準として複数のビデオフレームをフレーム単位で順次出力するときのビデオフレーム出力タイミングを調整させるタイミング調整ステップとを設けるようにする。

10

【0009】

エンコーダ側の基準クロックとデコーダ側のシステムタイムクロックとの間におけるクロック周波数のずれによって生じる時間差に応じ、複数のオーディオフレームをフレーム単位で順次出力するときのオーディオフレーム出力タイミングを基準として複数のビデオフレームをフレーム単位で順次出力するときのビデオフレーム出力タイミングを調整することにより、エンコーダ側とデコーダ側のクロック周波数の差を吸収し、オーディオフレーム出力タイミングにビデオフレーム出力タイミングを合わせてリップシンクさせることができる。

20

【0010】

さらに本発明においては、コンテンツ提供装置とコンテンツ受信装置を有するコンテンツ提供システムであって、コンテンツ提供装置は、エンコーダ側の基準クロックに基づくビデオタイムスタンプを付した複数の符号化ビデオフレームと、基準クロックに基づくオーディオタイムスタンプを付した複数の符号化オーディオフレームとを生成する符号化手段と、複数の符号化ビデオフレーム及び複数の符号化オーディオフレームをコンテンツ受信装置へ順次送信する送信手段とを具え、コンテンツ受信装置は、ビデオタイムスタンプが順次付された複数の符号化ビデオフレームと、オーディオタイムスタンプが順次付された複数の符号化オーディオフレームとをエンコーダ側のコンテンツ提供装置から受信して復号する復号手段と、復号手段によって符号化ビデオフレーム及び符号化オーディオフレームを復号した結果得られる複数のビデオフレーム及び複数のオーディオフレームを蓄積する記憶手段と、エンコーダ側の基準クロックのクロック周波数とデコーダ側のシステムタイムクロックのクロック周波数とのずれによって生じる時間差を算出する算出手段と、時間差に応じ、複数のオーディオフレームをフレーム単位で順次出力するときのオーディオフレーム出力タイミングを基準として複数のビデオフレームをフレーム単位で順次出力するときのビデオフレーム出力タイミングを調整するタイミング調整手段とを設けるようにする。

30

40

【0011】

エンコーダ側の基準クロックとデコーダ側のシステムタイムクロックとの間におけるクロック周波数のずれによって生じる時間差に応じ、複数のオーディオフレームをフレーム単位で順次出力するときのオーディオフレーム出力タイミングを基準として複数のビデオフレームをフレーム単位で順次出力するときのビデオフレーム出力タイミングを調整することにより、エンコーダ側とデコーダ側のクロック周波数の差を吸収し、オーディオフレーム出力タイミングにビデオフレーム出力タイミングを合わせてリップシンクさせることができる。

【発明の効果】

【0012】

50

上述のように本発明によれば、エンコーダ側の基準クロックとデコーダ側のシステムタイムクロックとの間におけるクロック周波数のずれによって生じる時間差に応じ、複数のオーディオフレームをフレーム単位で順次出力するときのオーディオフレーム出力タイミングを基準として複数のビデオフレームをフレーム単位で順次出力するときのビデオフレーム出力タイミングを調整することにより、エンコーダ側とデコーダ側のクロック周波数の差を吸収し、オーディオフレーム出力タイミングにビデオフレーム出力タイミングを合わせてリップシンクさせることができ、かくして視聴者であるユーザに違和感を感じさせることなく映像及び音声間のリップシンクを当該デコーダ側で確実に調整し得るコンテンツ受信装置、ビデオオーディオ出力タイミング制御方法及びコンテンツ提供システムを実現することができる。

10

【発明を実施するための最良の形態】

【0013】

以下、図面について、本発明の一実施の形態を詳述する。

【0014】

(1) コンテンツ提供システムの全体構成

図1において、1は全体として本発明のコンテンツ提供システムを示し、大きく分けてコンテンツ配信側となるコンテンツ提供装置2と、コンテンツ受信側となる第1のコンテンツ受信装置3及び第2のコンテンツ受信装置4とによって構成されている。

【0015】

コンテンツ提供システム1では、コンテンツ提供装置2及びWebサーバ14と第1のコンテンツ受信装置3とがインターネット5を介して相互に接続されており、当該Webサーバ14からインターネット5経由で取得したコンテンツ配信先であるURL(Uniform Resource Locator)や当該コンテンツに関するメタデータを第1のコンテンツ受信装置3におけるWebブラウザ15で解析し、そのメタデータやURLをストリーミングデコーダ9へ供給する。

20

【0016】

ストリーミングデコーダ9では、Webブラウザ15により解析したURLに基づいて当該コンテンツ提供装置2のストリーミングサーバ8へアクセスし、ユーザ所望のコンテンツに対する配信要求を行う。

【0017】

コンテンツ提供装置2は、エンコーダ7でユーザ所望のコンテンツに対応したコンテンツデータを予めエンコードし、その結果得られるエレメンタリストリームをストリーミングサーバ8でパケット化し、これをインターネット5を介して第1のコンテンツ受信装置3へ配信するようになされている。

30

【0018】

これによりコンテンツ提供システム1では、第1のコンテンツ受信装置3からの要求に応じたユーザ所望のコンテンツをコンテンツ提供装置2から配信するビデオオンデマンド(VOD)のようなプリエンコードドストリーミングを実現し得るようになされている。

【0019】

第1のコンテンツ受信装置3は、エレメンタリストリームをストリーミングデコーダ9でデコードすることにより元の映像及び音声を復元し、当該元の映像及び音声をモニタ10から出力するようになされている。

40

【0020】

またコンテンツ提供システム1では、第1のコンテンツ受信装置3と第2のコンテンツ受信装置4とが例えばIEEE(Institute of Electrical and Electronics Engineers)802.11a/b/g等の規格に準拠した無線LAN6で接続されており、当該第1のコンテンツ受信装置3が外部から供給された地上波デジタル、BS(Broadcast Satellite)/CS(Communication Satellite)デジタル又は地上波アナログ放送等のコンテンツあるいはDVD(Digital Versatile Disc)、VideoCDほか一般的なビデオカメラからのコンテンツをリアル

50

タイムストリーミングエンコーダ 11 でリアルタイムにエンコードした後に中継する形で第 2 のコンテンツ受信装置 4 へ無線送信し得るようになされている。

【0021】

因みに、第 1 のコンテンツ受信装置 3 と第 2 のコンテンツ受信装置 4 とは必ずしも無線 LAN 6 で接続されていなければならない訳ではなく、有線 LAN で接続されていても良い。

【0022】

第 2 のコンテンツ受信装置 4 は、第 1 のコンテンツ受信装置 3 から受信したコンテンツをリアルタイムストリーミングデコーダ 12 でデコードすることによりストリーミング再生を行い、その再生結果をモニタ 13 へ出力するようになされている。

10

【0023】

かくして第 1 のコンテンツ受信装置 3 及び第 2 のコンテンツ受信装置 4 の間では、外部から供給を受けたコンテンツを第 1 のコンテンツ受信装置 3 におけるリアルタイムストリーミングエンコーダ 11 でリアルタイムにエンコードして第 2 のコンテンツ受信装置 4 へ送信し、当該第 2 のコンテンツ受信装置 4 でストリーミング再生することにより、ライブストリーミングを実現し得るようになされている。

【0024】

(2) コンテンツ提供装置の構成

図 2 に示すようにコンテンツ提供装置 2 は、エンコーダ 7 及びストリーミングサーバ 8 によって構成されており、外部から取り込んだビデオ信号 VS 1 をビデオ入力部 21 を介してデジタル変換した後にビデオデータ VD 1 としてビデオエンコーダ 22 へ送出する。

20

【0025】

ビデオエンコーダ 22 は、ビデオデータ VD 1 を例えば MPEG 1/2/4 (Moving Picture Experts Group) の規格に準拠した所定の圧縮符号化方法あるいは種々の圧縮符号化方式で圧縮符号化し、その結果得られるビデオエレメンタリストリーム VES 1 をリングバッファでなるビデオ ES 蓄積部 23 へ送出する。

【0026】

ビデオ ES 蓄積部 23 は、ビデオエレメンタリストリーム VES 1 を一旦蓄積した後に、当該ビデオエレメンタリストリーム VES 1 をストリーミングサーバ 8 のパケット生成部 27 及びビデオフレームカウンタ 28 へ送出する。

30

【0027】

ビデオフレームカウンタ 28 では、ビデオエレメンタリストリーム VES 1 をフレーム周波数単位 (29.97 [Hz] あるいは 30 [Hz] あるいは 59.94 [Hz] あるいは 60 [Hz]) でカウントし、そのカウントアップ値を基準クロックに基づく 90 [KHz] 単位の値に変換し、32 ビット表現で各ビデオフレームに対するビデオタイムスタンプ VTS (VTS 1、VTS 2、VTS 3、...) としてパケット生成部 27 へ送出する。

【0028】

また、コンテンツ提供装置 2 は外部から取り込んだオーディオ信号 AS 1 をエンコーダ 7 のオーディオ入力部 24 を介してデジタル変換した後にオーディオデータ AD 1 としてオーディオエンコーダ 25 へ送出する。

40

【0029】

オーディオエンコーダ 25 は、オーディオデータ AD 1 を MPEG 1/2/4 オーディオの規格に準拠した所定の圧縮符号化方法あるいは種々の圧縮符号化方式で圧縮符号化し、その結果得られるオーディオエレメンタリストリーム AES 1 をリングバッファでなるオーディオ ES 蓄積部 26 へ送出する。

【0030】

オーディオ ES 蓄積部 26 は、オーディオエレメンタリストリーム AES 1 を一旦蓄積した後に、当該オーディオエレメンタリストリーム AES 1 をストリーミングサーバ 8 のパケット生成部 27 及びオーディオフレームカウンタ 29 へ送出する。

50

【0031】

オーディオフレームカウンタ29はビデオフレームカウンタ28と同様、オーディオフレームのカウントアップ値をビデオと共通の基準クロックに基づく90 [KHz]単位の値に変換し、各オーディオフレームに対するオーディオタイムスタンプATS (ATS1、ATS2、ATS3、...)として32ビット表現し、パケット生成部27へ送出する。

【0032】

パケット生成部27では、ビデオエレメンタリストリームVES1を所定データサイズのパケットに分割し、それぞれのパケットにビデオヘッダ情報を付加することによりビデオパケットを生成すると共に、オーディオエレメンタリストリームAES1を所定データサイズのパケットに分割し、それぞれのパケットにオーディオヘッダ情報を付加することによりオーディオパケットを生成する。

10

【0033】

ここで図3に示すようにオーディオパケット及びビデオパケットは、インターネット層におけるホスト間通信のIP (Internet Protocol)ヘッダ、トランスポート層における伝送制御用のTCP (Transmission Control Protocol)ヘッダ、リアルタイム・データ転送制御用のRTP (RealTime Transport Protocol)ヘッダ及びRTPペイロードからなり、RTPヘッダ内における4バイトのタイムスタンプ領域に上述のオーディオタイムスタンプATSやビデオタイムスタンプVTSが書き込まれるようになされている。

【0034】

そしてパケット生成部27 (図2)では、ビデオパケット及びビデオタイムスタンプVTSを基に所定バイト数からなる映像パケットデータを生成すると共に、オーディオパケット及びオーディオタイムスタンプATSを基に所定バイト数からなる音声パケットデータを生成し、これらを多重化することにより多重化データMXD1を生成した後パケットデータ蓄積部30へ送出する。

20

【0035】

パケットデータ蓄積部30は、多重化データMXD1を所定量蓄積すると、当該多重化データMXD1をインターネット5を介してRTP/TCP (RealTime Transport Protocol/Transmission Control Protocol)で第1のコンテンツ受信装置3へ送信するようになされている。

【0036】

(3) 第1のコンテンツ受信装置におけるストリーミングデコーダのモジュール構成

図4に示すように第1のコンテンツ受信装置3のストリーミングデコーダ9は、コンテンツ提供装置2からRTP/TCPで送信された多重化データMXD1を入力パケット蓄積部31に一旦蓄積した後、パケット分割部32へ送出する。

30

【0037】

ここで入力パケット蓄積部31は、インターネット5経由で送信されてくる多重化データMXD1が所定量のパケット分蓄積された時点で当該多重化データMXD1をパケット分割部32へ送出するようになされており、これにより後段のパケット分割部32で多重化データMXD1の処理が途切れることなく連続的に実行し得るようになされている。

【0038】

パケット分割部32は、多重化データMXD1を映像パケットデータVP1と音声パケットデータAP1に分割し、当該音声パケットデータAP1をリングバッファでなる入力オーディオバッファ33を介してオーディオフレーム単位でオーディオデコーダ35へ送出すると共に、映像パケットデータVP1をリングバッファでなる入力ビデオバッファ34を介してフレーム単位でビデオデコーダ36へ送出するようになされている。

40

【0039】

ここで入力オーディオバッファ33及び入力ビデオバッファ34においては、後段のオーディオデコーダ35及びビデオデコーダ36で1オーディオフレーム分の音声パケットデータAP1及び1ビデオフレーム分の映像パケットデータVP1を連続してデコードできるようにするまで蓄積するようになされており、そのため、いつの時点でも少なくとも

50

1 オーディオフレーム及び1ビデオフレーム分のデータをオーディオデコーダ35及びビデオデコーダ36へ瞬時に供給できるための容量を有する。

【0040】

なおパケット分割部32は、映像パケットデータVP1のビデオヘッダ情報及び音声パケットデータAP1のオーディオヘッダ情報を解析することによりビデオタイムスタンプVTS及びオーディオタイムスタンプATSを認識し得るようになされており、当該ビデオタイムスタンプVTS及び当該オーディオタイムスタンプATSをレンダラ-37のタイミングコントロール回路37Aへ送出する。

【0041】

オーディオデコーダ35は、音声パケットデータAP1をオーディオフレーム単位でデコードすることにより圧縮符号化前のオーディオフレームAF1を復元し、順次レンダラ-37へ送出する。

10

【0042】

ビデオデコーダ36は、映像パケットデータVP1をビデオフレーム単位でデコードすることにより圧縮符号化前のビデオフレームVF1を復元し、順次レンダラ-37へ送出する。

【0043】

ところでストリーミングデコーダ9においては、Webブラウザ15からシステムコントローラ50に対してコンテンツのメタデータMDが供給されており、コンテンツ判別手段としての当該システムコントローラ50では当該メタデータMDに基づいて当該コンテンツの種類がオーディオ及びビデオからなるものであるか、ビデオだけからなるものであるか、或いはオーディオだけからなるものであるかを判別し、そのコンテンツ種類判別結果CHをレンダラ-37へ送出する。

20

【0044】

レンダラ-37は、オーディオフレームAF1をリングバッファでなる出力オーディオバッファ38へ一時的に格納し、また同様にビデオフレームVF1をリングバッファでなる出力ビデオバッファ39に一時的に格納する。

【0045】

そしてレンダラ-37は、タイミングコントロール回路37Aによってモニタ10へ出力すべきビデオフレームVF1の映像とオーディオフレームAF1の音声とをリップシンクさせるべく、システムコントローラ50からのコンテンツ種類判別結果CHと、オーディオタイムスタンプATSやビデオタイムスタンプVTSとに基づいて最終的な出力タイミングを調整した後、その出力タイミングで出力ビデオバッファ39、出力オーディオバッファ38からビデオフレームVF1、オーディオフレームAF1を順次出力するようになされている。

30

【0046】

(4) プリエンコードストリーミングにおけるデコーダ側でのリップシンク調整処理
(4-1) プリエンコードストリーミングにおけるビデオフレーム及びオーディオフレームの出力タイミング調整方法

図5に示すようにレンダラ-37のタイミングコントロール回路37Aでは、パケット分割部32から送られたビデオタイムスタンプVTS(VTS1、VTS2、VTS3、...、VTSn)及びオーディオタイムスタンプATS(ATS1、ATS2、ATS3、...、ATSn)をバッファ42及び43にそれぞれ一時的に格納した後、コンパレータ回路46へ送出する。

40

【0047】

またタイミングコントロール回路37Aは、そのコンテンツにおける最初のビデオタイムスタンプVTS1及びオーディオタイムスタンプATS1だけをサブトラクタ回路44及び45にもそれぞれ送出する。

【0048】

サブトラクタ回路44及び45は、当該最初のビデオタイムスタンプVTS1及びオー

50

ディオタイムスタンプ A T S 1 の値を所定時間分だけ引き戻し、これらをプリセット用ビデオタイムスタンプ V T S p 及びプリセット用オーディオタイムスタンプ A T S p として S T C 回路 4 1 へ送出する。

【 0 0 4 9 】

S T C 回路 4 1 では、システムタイムクロック s t c の値をプリセット用ビデオタイムスタンプ V T S p、プリセット用オーディオタイムスタンプ A T S p の順番で決められたプリセットシーケンスに従ってプリセットする、すなわち当該システムタイムクロック s t c の値をプリセット用ビデオタイムスタンプ V T S p、プリセット用オーディオタイムスタンプ A T S p の順番でアジャストする（置き換える）ようになされている。

【 0 0 5 0 】

ここで S T C 回路 4 1 では、当該最初のビデオタイムスタンプ V T S 1 及びオーディオタイムスタンプ A T S 1 の値を所定時間分だけ引き戻したプリセット用ビデオタイムスタンプ V T S p 及びプリセット用オーディオタイムスタンプ A T S p を用いてシステムタイムクロック s t c の値をプリセットするため、バッファ 4 2 及び 4 3 を介して当該最初のビデオタイムスタンプ V T S 1 及びオーディオタイムスタンプ A T S 1 がコンパレータ回路 4 6 に到達したとき、S T C 回路 4 1 からコンパレータ回路 4 6 へ供給されるプリセット後のシステムタイムクロック s t c の値が当該ビデオタイムスタンプ V T S 1 及びオーディオタイムスタンプ A T S 1 よりも前の時刻を示すようになされている。

【 0 0 5 1 】

これによりタイミングコントロール回路 3 7 A のコンパレータ回路 4 6 では、プリセット後のシステムタイムクロック s t c の値が最初のビデオタイムスタンプ V T S 1 及びオーディオタイムスタンプ A T S 1 に対して既に経過しているといったことが無くなるため、当該最初のビデオタイムスタンプ V T S 1 及びオーディオタイムスタンプ A T S 1 に対応したビデオフレーム V f 1 及びオーディオフレーム A f 1 についても確実に出力し得るようになされている。

【 0 0 5 2 】

實際上、図 6 (A) 及び (B) に示すように、コンテンツの種類がオーディオ及びビデオからなるものである場合、システムタイムクロック s t c の値をプリセット用ビデオタイムスタンプ V T S p、プリセット用オーディオタイムスタンプ A T S p の順番で決められたプリセットシーケンスに従ってプリセットすると、プリセット用ビデオタイムスタンプ V T S p でシステムタイムクロック s t c の値をプリセットした後に必ずプリセット用オーディオタイムスタンプ A T S p で先程のプリセット値が更新されることを意味する。

【 0 0 5 3 】

このときコンパレータ回路 4 6 は、プリセット用オーディオタイムスタンプ A T S p でプリセット値が更新された後のシステムタイムクロック s t c を基準にしてビデオタイムスタンプ V T S と比較することにより、プリセット後のシステムタイムクロック s t c の値とエンコーダ側のコンテンツ提供装置 2 で付けられたビデオタイムスタンプ V T S との時間差を算出するようになされている。

【 0 0 5 4 】

一方、コンテンツの種類がオーディオだけからなるものである場合にはプリセット用ビデオタイムスタンプ V T S p がタイミングコントロール回路 3 7 A に送られてくることはないので、システムタイムクロック s t c の値をプリセット用ビデオタイムスタンプ V T S p、プリセット用オーディオタイムスタンプ A T S p の順番で決められたプリセットシーケンスに従えば、当然プリセット用オーディオタイムスタンプ A T S p でシステムタイムクロック s t c の値がプリセットされることを意味する。

【 0 0 5 5 】

同様に、コンテンツの種類がビデオだけからなるものである場合にはプリセット用オーディオタイムスタンプ A T S p がタイミングコントロール回路 3 7 A に送られてくることはないので、システムタイムクロック s t c の値をプリセット用ビデオタイムスタンプ V T S p、プリセット用オーディオタイムスタンプ A T S p の順番で決められたプリセット

10

20

30

40

50

シーケンスに従えば、当然プリセット用ビデオタイムスタンプV T S pでシステムタイムクロックs t cの値がプリセットされることを意味する。

【0056】

これは、コンテンツの種類がオーディオだけでなる場合、若しくはビデオだけでなる場合だけであり、映像及び音声のリップシンクを調整する必要は特にないため、プリセット用オーディオタイムスタンプA T S pでプリセットされた後のシステムタイムクロックs t cの値とオーディオタイムスタンプA T Sとが一致したときにオーディオフィームA F 1を出力すればよく、またプリセット用ビデオタイムスタンプV T S pでプリセットされた後のシステムタイムクロックs t cの値とビデオタイムスタンプV T Sとが一致したときにビデオフレームV F 1を出力すればよい。

10

【0057】

實際上、レンダラー37のタイミングコントロール回路37Aでは、例えばコンテンツの種類がオーディオ及びビデオからなるものである場合、図7に示すように例えばオーディオデコーダ35でデコードした後のオーディオフィームA F 1 (A f 1、A f 2、A f 3、...) をモニタ10へ順次出力する時点T a 1、T a 2、T a 3、...、のタイミングでは、クリスタルオシレータ回路40 (図4) 及びS T C回路41を介して供給されるシステムタイムクロックs t cの値をプリセット用ビデオタイムスタンプV T S p、プリセット用オーディオタイムスタンプA T S pの順番でプリセットすることにより、最終的にシステムタイムクロックs t cの値をプリセット用オーディオタイムスタンプA T S p 1、A T S p 2、A T S p 3、...と一致させる。

20

【0058】

このことは、再生中に音声途切り音飛びがあるとユーザにとって非常に目立つので、レンダラー37のタイミングコントロール回路37AではオーディオフィームA F 1 (A f 1、A f 2、A f 3、...) をリップシンク調整処理の基準として用い、当該オーディオフィームA F 1 (A f 1、A f 2、A f 3、...) の出力に合わせてビデオフレームV F 1 (V f 1、V f 2、V f 3、...) の出力タイミングを調整する必要があるからである。

【0059】

またレンダラー37のタイミングコントロール回路37Aは、オーディオフィームA F 1 (A f 1、A f 2、A f 3、...) の出力タイミング (時点T a 1、T a 2、T a 3、...) が決まると、ビデオフレームV F 1 (V f 1、V f 2、V f 3、...) をシステムタイムクロックs t cに基づく30 [Hz]のフレーム周波数で出力する任意の時点T v 1、T v 2、T v 3、...において、プリセット後のシステムタイムクロックs t cのカウント値と、ビデオフレームV F 1 (V f 1、V f 2、V f 3、...) に付されているビデオタイムスタンプV T S (V T S 1、V T S 2、V T S 3、...) とをコンパレータ回路46でそれぞれ比較する。

30

【0060】

コンパレータ回路46では、プリセット後のシステムタイムクロックs t cのカウント値と、ビデオタイムスタンプV T S (V T S 1、V T S 2、V T S 3、...) とが一致したときに出力ビデオバッファ39からビデオフレームV F 1 (V f 1、V f 2、V f 3、...) をモニタ10へ出力させるようになされている。

40

【0061】

ところでコンパレータ回路46は、プリセット後のシステムタイムクロックs t cのカウント値と、バッファ42から送られるビデオタイムスタンプV T S (V T S 1、V T S 2、V T S 3、...) とを比較した結果、プリセット後のシステムタイムクロックs t cのカウント値とビデオタイムスタンプV T S (V T S 1、V T S 2、V T S 3、...) との差分値D 1 (時間差) が所定の時間を表す閾値T H以下であれば、ユーザにとっては映像と音声とが一致していないとは認識し得ないレベルなので、タイミングコントロール回路37Aはプリセット後のシステムタイムクロックs t cのカウント値とビデオタイムスタンプV T S (V T S 1、V T S 2、V T S 3、...) とが一致したときにビデオフレ

50

△VF1 (Vf 1、Vf 2、Vf 3、...) をそのままモニタ 10 に出力すればよい。

【0062】

それ以外の場合、例えば時点 Tv 2 のタイミングにおいて、プリセット後のシステムタイムクロック stc のカウント値とビデオタイムスタンプ VTS 2 との差分値 D 1 が所定の閾値 TH よりも大きく、かつ映像が音声よりも遅れている場合には、エンコーダ側のクロック周波数とデコーダ側のクロック周波数とのずれが原因で音声に映像が追いついていない状態であるため、レンダラ 37 のタイミングコントロール回路 37 A では GOP (Group Of Picture) を構成している例えば B ピクチャに相当するビデオフレーム Vf 3 (図示せず) をデコードすることなくスキップし、次のビデオフレーム Vf 4 を出力するようになされている。

10

【0063】

この場合、レンダラ 37 は出力ビデオバッファ 39 に格納されている「P」ピクチャについては、ビデオデコーダ 36 で次のピクチャをデコードする際の参照フレームとなるためスキップせず、次のピクチャを生成する際の参照フレームとならない非参照フレームである「B」ピクチャをスキップすることにより、画質劣化を未然に防ぎながらリップシンクさせるようになされている。

【0064】

ところでレンダラ 37 では、仮にスキップすべき「B」ピクチャが出力ビデオバッファ 39 に存在せず、「I」ピクチャや「P」ピクチャばかりであった場合には、当該「B」ピクチャをスキップすることはできないため、音声に映像を追い付かせることができなくなってしまう。

20

【0065】

そこでレンダラ 37 では、スキップすべき「B」ピクチャが出力ビデオバッファ 39 に存在しないときには、図 8 に示すようにモニタ 10 のモニタ出力タイミングが例えば 60 [Hz] であり、出力ビデオバッファ 39 から出力すべきビデオフレーム VF 1 のピクチャリフレッシュタイミングが 30 [Hz] であることを利用し、当該ピクチャリフレッシュタイミングを短縮するようになされている。

【0066】

具体的にはレンダラ 37 は、プリセット用オーディオタイムスタンプ ATS p でプリセットした後のシステムタイムクロック stc のカウント値とビデオタイムスタンプ VTS の差分値 D 1 が 16.666... [msec] を超えるとき、すなわち音声の出力タイミングに対してモニタ出力タイミングが 1 フレーム分以上遅れているときは、1 フレーム分のビデオフレーム VF 1 をスキップする代わりにピクチャリフレッシュタイミングを 30 [Hz] から 60 [Hz] に変更して次の N + 1 番目のピクチャを出力するようになされている。

30

【0067】

つまりレンダラ 37 は、当該スキップによる画質劣化の影響を受ける「I」ピクチャや「P」ピクチャについてはピクチャリフレッシュ間隔を 1 / 30 秒から 1 / 60 秒に短縮することにより、「I」ピクチャや「P」ピクチャをスキップすることによる画質劣化を生じさせることなく音声に映像を追い付かせることができるようになされている。

【0068】

これに対してレンダラ 37 のタイミングコントロール回路 37 A は、時点 Tv 2 のタイミングにおいて、プリセット後のシステムタイムクロック stc のカウント値と例えばビデオタイムスタンプ VTS 2 との差分値 D 1 が所定の閾値 TH よりも大きく、かつ音声映像が映像よりも遅れている場合には、エンコーダ側のクロック周波数とデコーダ側のクロック周波数とのずれが原因で映像に音声追いついていない状態であるため、現在出力中のビデオフレーム Vf 2 を繰り返しリピートして出力するようになされている。

40

【0069】

一方、レンダラ 37 のタイミングコントロール回路 37 A では、例えばコンテンツの種類がビデオだけからなるものである場合、ビデオデコーダ 36 でデコードした後のビデオフレーム VF 1 (Vf 1、Vf 2、Vf 3、...) をモニタ 10 へ順次出力する時点 T

50

v 1、T v 2、T v 3、...、のタイミングでは、プリセット用ビデオタイムスタンプ V T S p でプリセットされたシステムタイムクロック s t c のカウント値とビデオタイムスタンプ V T S が一致したタイミングでビデオフレーム V F 1 (V f 1、V f 2、V f 3、...) をモニタ 1 0 に出力すればよい。

【 0 0 7 0 】

同様に、レンダラー 3 7 のタイミングコントロール回路 3 7 A では、例えばコンテンツの種類がオーディオだけからなるものである場合、オーディオデコーダ 3 5 でデコードした後のオーディオフレーム A F 1 (A f 1、A f 2、A f 3、...) をモニタ 1 0 へ順次出力する時点 T a 1、T a 2、T a 3、...、のタイミングでは、プリセット用オーディオタイムスタンプ A T S p でプリセットされたシステムタイムクロック s t c のカウント値とオーディオタイムスタンプ A T S が一致したタイミングでオーディオフレーム A F 1 (A f 1、A f 2、A f 3、...) をモニタ 1 0 のスピーカから出力すればよい。

10

【 0 0 7 1 】

(4 - 2) プリエンコードストリーミングにおけるリップシンク調整処理手順

上述のようにストリーミングデコーダ 9 におけるレンダラー 3 7 のタイミングコントロール回路 3 7 A がオーディオフレーム A F 1 (A f 1、A f 2、A f 3、...) を基準にしてビデオフレーム V F 1 (V f 1、V f 2、V f 3、...) の出力タイミングを調整することにより、映像と音声とをリップシンクさせる出力タイミング調整方法についてまとめると、次の図 9 のフローチャートに示すように、レンダラー 3 7 のタイミングコントロール回路 3 7 A は、ルーチン R T 1 の開始ステップから入って、次のステップ S P 1 へ移

20

【 0 0 7 2 】

ステップ S P 1 においてレンダラー 3 7 は、システムタイムクロック s t c の値をプリセット用ビデオタイムスタンプ V T S p、プリセット用オーディオタイムスタンプ A T S p の順番で決められたプリセットシーケンスに従ってプリセットし、次のステップ S P 2 へ移る。

【 0 0 7 3 】

ここでレンダラー 3 7 は、コンテンツの種類がオーディオ及びビデオでなるものであるときにはプリセット用ビデオタイムスタンプ V T S p でシステムタイムクロック s t c の値をプリセットした後に必ずプリセット用オーディオタイムスタンプ A T S p で先程のプリセット値を更新し、次のステップ S P 2 へ移る。

30

【 0 0 7 4 】

この場合、オーディオフレーム A F 1 (A f 1、A f 2、A f 3、...) をモニタ 1 0 へ出力する時点 T a 1、T a 2、T a 3、... のタイミングで (図 7)、システムタイムクロック s t c の値とプリセット用オーディオタイムスタンプ A T S p (A T S p 1、A T S p 2、A T S p 3、...) とが一致することになる。

【 0 0 7 5 】

またレンダラー 3 7 は、コンテンツの種類がビデオだけからなるものである場合には、プリセット用オーディオタイムスタンプ A T S p は存在しないので、プリセット用ビデオタイムスタンプ V T S p でシステムタイムクロック s t c の値をプリセットして所定時間経過したときに次のステップ S P 2 へ移る。

40

【 0 0 7 6 】

さらにレンダラー 3 7 は、コンテンツの種類がオーディオだけからなるものである場合には、プリセット用ビデオタイムスタンプ V T S p は存在しないので、プリセット用ビデオタイムスタンプ V T S p を待つことなくプリセット用オーディオタイムスタンプ A T S p が到達した時点でシステムタイムクロック s t c の値をプリセットした後に次のステップ S P 2 へ移る。

【 0 0 7 7 】

ステップ S P 2 においてレンダラー 3 7 は、システムコントローラ 5 0 から供給されるコンテンツ種類判別結果 C H に基づいて当該コンテンツがビデオのみでなるものか否かを

50

判定し、肯定結果が得られると次のステップ S P 3 へ移る。

【 0 0 7 8 】

ステップ S P 3 においてレンダーラ- 3 7 は、当該コンテンツがビデオのみでなるため、プリセット用ビデオタイムスタンプ V T S p でプリセットしたシステムタイムクロック s t c のカウント値とビデオタイムスタンプ V T S とが一致したときにビデオフレーム V F 1 (V f 1、V f 2、V f 3、...) をモニタ 1 0 へ出力し、次のステップ S P 1 2 へ移って処理を終了する。

【 0 0 7 9 】

これに対してステップ S P 2 で否定結果が得られると、このことはコンテンツの種類がビデオのみでなるものではなく、オーディオ及びビデオでなるものか、オーディオのみでなるものかの何れかであることを表しており、このときレンダーラ- 3 7 は次のステップ S P 4 へ移る。

【 0 0 8 0 】

ステップ S P 4 においてレンダーラ- 3 7 は、コンテンツ種類判別結果 C H に基づいて当該コンテンツがオーディオのみでなるものか否かを判定し、肯定結果が得られると次のステップ S P 3 へ移る。

【 0 0 8 1 】

ステップ S P 3 においてレンダーラ- 3 7 は、当該コンテンツがオーディオのみでなるため、プリセット用オーディオタイムスタンプ A T S p でプリセットしたシステムタイムクロック s t c のカウント値とオーディオタイムスタンプ A T S とが一致したときにオーディオフレーム A F 1 (A f 1、A f 2、A f 3、...) をモニタ 1 0 のスピーカから出力し、次のステップ S P 1 2 へ移って処理を終了する。

【 0 0 8 2 】

これに対してステップ S P 4 で否定結果が得られると、このことはコンテンツの種類がオーディオ及びビデオからなるものであることを表しており、このときレンダーラ- 3 7 は次のステップ S P 5 へ移る。

【 0 0 8 3 】

ステップ S P 5 においてレンダーラ- 3 7 は、コンテンツの種類がオーディオ及びビデオからなるものであるため、最終的にプリセット用オーディオタイムスタンプ A T S p でプリセットされたシステムタイムクロック s t c のカウント値と、時点 T v 1、T v 2、T v 3、... のタイミングで出力すべきビデオフレーム V F 1 (V f 1、V f 2、V f 3、...) のタイムスタンプ V T S (V T S 1、V T S 2、V T S 3、...) との差分値 D 1 (= s t c - V T S) を算出し、次のステップ S P 6 へ移る。

【 0 0 8 4 】

ステップ S P 6 においてレンダーラ- 3 7 は、ステップ S P 7 で算出した差分値 D 1 (絶対値) が所定の閾値 T H よりも大きいかなかを判定する。ここで否定結果が得られると、このことは、差分値 D 1 が、映像及び音声を見て聞いたユーザにとって当該映像と当該音声との間にずれが生じているとは判断し得ない程度の時間 (例えば 1 0 0 [msec]) 以下であることを表しており、このときレンダーラ- 3 7 は次のステップ S P 3 へ移る。

【 0 0 8 5 】

ステップ S P 3 においてレンダーラ- 3 7 は、映像と音声がずれていると判断し得ない程度の時間差しかないので、この場合は当該ビデオフレーム V F 1 をそのままモニタ 1 0 へ出力し、またオーディオフレーム A F 1 についても原則的にそのままモニタ 1 0 へ出力し、次のステップ S P 1 2 へ移って処理を終了する。

【 0 0 8 6 】

これに対してステップ S P 6 で肯定結果が得られると、このことは差分値 D 1 が所定の閾値 T H よりも大きい、すなわち映像及び音声を見て聞いたユーザにとって当該映像と当該音声との間にずれが生じていると判断し得る程度であることを表しており、このときレンダーラ- 3 7 は次のステップ S P 7 へ移る。

【 0 0 8 7 】

10

20

30

40

50

ステップ S P 7 においてレンダーラ-37 は、映像が音声よりも遅れているか否かをオーディオタイムスタンプ A T S 及びビデオタイムスタンプ V T S に基づいて判定し、否定結果が得られると次のステップ S P 8 へ移る。

【0088】

ステップ S P 8 においてレンダーラ-37 は、映像の方が音声よりも進んでいるので、当該映像に音声が進んでいくように現在出力中のピクチャを構成しているビデオフレーム V F 1 を繰り返しリピート出力した後、次のステップ S P 12 へ移って処理を終了する。

【0089】

これに対してステップ S P 7 で肯定結果が得られると、このことは映像が音声よりも遅れていることを表しており、このときレンダーラ-37 は次のステップ S P 9 へ移り、出力ビデオバッファ 39 にスキップ対象の「B」ピクチャが存在するか否かを判定し、肯定結果が得られると次のステップ S P 10 へ移る。

10

【0090】

ステップ S P 10 においてレンダーラ-37 は、音声に対する映像の遅れを取り戻すべく B ピクチャ（この場合、ビデオフレーム V f 3）をデコードせずにスキップして出力することにより、音声に対する映像の遅れを取り戻してリップシンクさせることができ、次のステップ S P 12 へ移って処理を終了する。

【0091】

一方、ステップ S P 9 で否定結果が得られると、このことは出力ビデオバッファ 39 にスキップ対象の「B」ピクチャが存在せず、「B」ピクチャをスキップすることができないことを表しており、このときレンダーラ-37 は次のステップ S P 11 へ移る。

20

【0092】

ステップ S P 11 においてレンダーラ-37 は、図 8 に示したように、モニタ 10 のモニタ出力タイミングが 60 [Hz] であるのに対し、ビデオフレーム V F 1 のピクチャリフレッシュタイミングが 30 [Hz] であることを利用し、当該ピクチャリフレッシュタイミングをモニタ 10 のモニタ出力タイミングに合わせて短縮することにより、ピクチャをスキップすることによる画質劣化を生じさせずに映像を音声に追い付かせ、次のステップ S P 12 へ移って処理を終了する。

【0093】

(5) 第 1 のコンテンツ受信装置におけるリアルタイムストリーミングエンコーダの回路構成

30

第 1 のコンテンツ受信装置 3（図 1）は、外部から供給された地上波デジタル、B S / C S デジタル又は地上波アナログ放送等のコンテンツあるいは D V D、VideoCD ほか一般的なビデオカメラからのコンテンツをリアルタイムストリーミングエンコーダ 11 によってリアルタイムにエンコードした後に第 2 のコンテンツ受信装置 4 へ中継する形で無線送信することによりコンテンツ提供側にもなり得るようになされている。

【0094】

その第 1 のコンテンツ受信装置 3 におけるリアルタイムストリーミングエンコーダ 11 の回路構成について図 10 を用いて説明する。リアルタイムストリーミングエンコーダ 11 は、外部から供給されたコンテンツを構成するビデオ信号 V S 2 及びオーディオ信号 A S 2 をビデオ入力部 51 及びオーディオ入力部 53 を介してデジタル変換し、これをビデオデータ V D 2 及びオーディオデータ A D 2 としてビデオエンコーダ 52 及びオーディオエンコーダ 54 へ送出する。

40

【0095】

ビデオエンコーダ 52 は、ビデオデータ V D 2 を例えば M P E G 1/2/4 の規格に準拠した所定の圧縮符号化方法あるいは種々の圧縮符号化方式で圧縮符号化し、その結果得られるビデオエレメンタリストリーム V E S 2 をパケット生成部 56 及びビデオフレームカウンタ 57 へ送出する。

【0096】

ビデオフレームカウンタ 57 では、ビデオエレメンタリストリーム V E S 2 をフレーム

50

周波数単位(29.97 [Hz]あるいは30 [Hz]あるいは59.94 [Hz]あるいは60 [Hz])でカウントし、そのカウントアップ値を基準クロックに基づく90 [KHz]単位の値に変換し、32ビット表現で各ビデオフレームに対するビデオタイムスタンプVTS(VTS1、VTS2、VTS3、...)としてパケット生成部56へ送出する。

【0097】

オーディオエンコーダ54は、オーディオデータAD2をMPEG1/2/4オーディオの規格に準拠した所定の圧縮符号化方法あるいは種々の圧縮符号化方式で圧縮符号化し、その結果得られるオーディオエレメンタリストリームAES2をパケット生成部56及びオーディオフレームカウンタ58へ送出する。

【0098】

オーディオフレームカウンタ58はビデオフレームカウンタ57と同様、オーディオフレームのカウントアップ値と共通の基準クロックに基づく90 [KHz]単位の値に変換し、オーディオタイムスタンプATS(ATS1、ATS2、ATS3、...)として32ビット表現し、パケット生成部56へ送出する。

【0099】

パケット生成部56では、ビデオエレメンタリストリームVES2を所定データサイズのパケットに分割し、それぞれのパケットにビデオヘッダ情報を付加することによりビデオパケットを生成すると共に、オーディオエレメンタリストリームAES2を所定データサイズのパケットに分割し、それぞれのパケットにオーディオヘッダ情報を付加することによりオーディオパケットを生成する。

【0100】

ここで図11に示すようにRTCP(Real Time Control Protocol)packetの前段に付加されているコントロールパケットは、インターネット層におけるホスト間通信用のIP(Internet Protocol)ヘッダ、ユーザ・データグラム・データ転送用のUDP(User Datagram Protocol)ヘッダ、リアルタイム・データ転送制御用のRTCP(Real Time Control Protocol)パケットセンドリポート及びRTCPパケットからなり、RTCPパケットセンドリポート内のセンド情報内にある4バイトのRTPタイムスタンプ領域にPCR(Program Clock Reference)値としてエンコーダ側におけるシステムタイムクロック値のスナッチショット情報が書き込まれるようになされていて、デコーダ側のクロックリカバリ用にPCR回路61から送出される。

【0101】

そしてパケット生成部56では、ビデオパケット及びビデオタイムスタンプVTSに基づいて所定バイト数からなる映像パケットデータを生成すると共に、オーディオパケット及びオーディオタイムスタンプATSに基づいて所定バイト数からなる音声パケットデータを生成し、これらを多重化することにより多重化データMXD2を生成した後パケットデータ蓄積部59へ送出する。

【0102】

パケットデータ蓄積部59は、多重化データMXD2を所定量蓄積すると、当該多重化データMXD2を無線LAN6を介してRTP/RTCPで第2のコンテンツ受信装置4へ送信するようになされている。

【0103】

ところでリアルタイムストリーミングエンコーダ11は、ビデオ入力部51でデジタル変換したビデオデータVD2をPLL(Phase-Locked Loop)回路55にも供給する。PLL回路55は、ビデオデータVD2に基づいて当該ビデオデータVD2のクロック周波数にSTC回路60を同期させると共に、ビデオエンコーダ52、オーディオ入力部53及びオーディオエンコーダ54についてもビデオデータVD2のクロック周波数と同期させるようになされている。

【0104】

これによりリアルタイムストリーミングエンコーダ11は、PLL回路55を介してビデオデータVD2に対する圧縮符号化処理とオーディオデータAD2に対する圧縮符号化

10

20

30

40

50

処理とをビデオデータV D 2のクロック周波数と同期したタイミングで実行し得ると共に、P C R (Program Clock Reference)回路6 1を介してビデオデータV D 2のクロック周波数に同期したクロックリファレンスp c rを第2のコンテンツ受信装置4におけるリアルタイムストリーミングデコーダ1 2へ送信し得るようになされている。

【0 1 0 5】

このときP C R回路6 1は、クロックリファレンスp c rをR T Pプロトコルの下位層に位置しリアルタイム性が要求されるU D P (User Datagram Protocol)で第2のコンテンツ受信装置4のリアルタイムストリーミングデコーダ1 2へ送信するようになされており、これにより高速性を確保してリアルタイム性の必要とされるライブストリーミングにも対応し得るようになされている。

10

【0 1 0 6】

(6)第2のコンテンツ受信装置におけるリアルタイムストリーミングデコーダの回路構成

図1 2に示すように第2のコンテンツ受信装置4におけるリアルタイムストリーミングデコーダ1 2は、第1のコンテンツ受信装置3のリアルタイムストリーミングエンコーダ1 1から送信された多重化データM X D 2を入力パケット蓄積部7 1に一旦蓄積した後、パケット分割部7 2へ送出する。

【0 1 0 7】

パケット分割部7 2は、多重化データM X D 2を映像パケットデータV P 2と音声パケットデータA P 2に分割し、当該音声パケットデータA P 2をリングバッファでなる入力オーディオバッファ7 3を介してオーディオフィーム単位でオーディオデコーダ7 4へ送出すると共に、映像パケットデータV P 2をリングバッファでなる入力ビデオバッファ7 5を介してフレーム単位でビデオデコーダ7 6へ送出するようになされている。

20

【0 1 0 8】

ここで入力オーディオバッファ7 3及び入力ビデオバッファ7 5においても、後段のオーディオデコーダ7 4及びビデオデコーダ7 6で1オーディオフィーム及び1ビデオフレーム分の音声パケットデータA P 2及び映像パケットデータV P 2を連続してデコードできるようにするまで蓄積するようになされており、そのため少なくとも1オーディオフィーム及び1ビデオフレーム分のデータ容量があればよい。

【0 1 0 9】

なおパケット分割部7 2は、映像パケットデータV P 2のビデオヘッダ情報及び音声パケットデータA P 2のオーディオヘッダ情報を解析することによりオーディオタイムスタンプA T S及びビデオタイムスタンプV T Sを認識し得るようになされており、当該オーディオタイムスタンプA T S及び当該ビデオタイムスタンプV T Sをレンダーラ-7 7へ送出する。

30

【0 1 1 0】

オーディオデコーダ7 4は、音声パケットデータA P 2をオーディオフィーム単位でデコードすることにより圧縮符号化前のオーディオフィームA F 2を復元し、順次レンダーラ-7 7へ送出する。

【0 1 1 1】

ビデオデコーダ7 6は、映像パケットデータV P 2をビデオフレーム単位でデコードすることにより圧縮符号化前のビデオフレームV F 2を復元し、順次レンダーラ-7 7へ送出する。

40

【0 1 1 2】

レンダーラ-7 7は、オーディオフィームA F 2をリングバッファでなる出力オーディオバッファ7 8へ一時的に格納し、また同様にビデオフレームV F 2をリングバッファでなる出力ビデオバッファ7 9に一時的に格納する。

【0 1 1 3】

そしてレンダーラ-7 7は、モニタ1 3へ出力するビデオフレームV F 2の映像とオーディオフィームA F 2の音声とをリップシンクさせるべくオーディオタイムスタンプA T S

50

及びビデオタイムスタンプV T Sに基づいて最終的な出力タイミングを調整した後、その出力タイミングで出力オーディオバッファ7 8及び出力ビデオバッファ7 9からオーディオフレームA F 2及びビデオフレームV F 2をモニタ1 3へ順次出力するようになされている。

【0 1 1 4】

ところでリアルタイムストリーミングデコーダ1 2は、第1のコンテンツ受信装置3におけるリアルタイムストリーミングエンコーダ1 1のP C R回路6 1からU D Pで送信されるクロックリファレンスp c rを受信して減算回路8 1に入力する。

【0 1 1 5】

減算回路8 1は、クロックリファレンスp c rとS T C回路8 4から供給されるシステムタイムクロックs t cとの差を算出し、これをフィルタ8 2、電圧制御型クリスタルオシレータ回路8 3及びS T C回路8 4を順次介して減算回路8 1にフィードバックすることによりP L L (Phase Locked Loop)を形成し、リアルタイムストリーミングエンコーダ1 1のクロックリファレンスp c rに次第に収束させ、最終的には当該クロックリファレンスp c rによりリアルタイムストリーミングエンコーダ1 1と同期したシステムタイムクロックs t cをレンダラー7 7へ供給するようになされている。

【0 1 1 6】

これによりレンダラー7 7は、第1のコンテンツ受信装置3におけるリアルタイムストリーミングエンコーダ1 1でビデオデータV D 2及びオーディオデータA D 2を圧縮符号化したり、ビデオタイムスタンプV T S及びオーディオタイムスタンプA T Sをカウントする際のクロック周波数と同期したシステムタイムクロックs t cを基準にして、ビデオフレームV F 2及びオーディオフレームA F 2の出力タイミングを調整し得るようになされている。

【0 1 1 7】

実際上レンダラー7 7は、オーディオフレームA F 2に関してはリングバッファでなる出力オーディオバッファ7 8へ一時的に格納すると共に、ビデオフレームV F 2に関してはリングバッファでなる出力ビデオバッファ7 9に一時的に格納し、映像と音声とをリップシンクさせた状態で出力するべく、リアルタイムストリーミングエンコーダ1 1のP C R回路6 1から供給されるクロックリファレンスp c rによりエンコーダ側と同期したシステムタイムクロックs t c及びオーディオタイムスタンプA T S、ビデオタイムスタンプV T Sに基づいて出力タイミングを調整するようになされている。

【0 1 1 8】

(7) ライブストリーミングにおけるデコーダ側でのリップシンク調整処理

(7-1) ライブストリーミングにおけるビデオフレーム及びオーディオフレームの出力タイミング調整方法

図1 3に示すように、この場合レンダラー7 7は、リアルタイムストリーミングエンコーダ1 1のP C R回路6 1から所定周期で供給されてくるクロックリファレンスp c rの値に、システムタイムクロックs t cのクロック周波数をP L Lでロックさせたうえで、当該システムタイムクロックs t cの基で同期されたモニタ1 3を通してオーディオタイムスタンプA T S及びビデオタイムスタンプV T Sに従いオーディオフレームA F 2及びビデオフレームV F 2の出力をコントロールする。

【0 1 1 9】

すなわちレンダラー7 7は、クロックリファレンスp c rの値にシステムタイムクロックs t cのクロック周波数を同期した状態で、システムタイムクロックs t c及びオーディオタイムスタンプA T S (A T S 1、A T S 2、A T S 3、... ..)に従ってオーディオフレームA F 2 (A f 1、A f 2、A f 3、... ..)をモニタ1 3へ順次出力する。

【0 1 2 0】

ここで、クロックリファレンスp c rの値とシステムタイムクロックs t cのクロック周波数とは前述のように同期関係を維持しているため、システムタイムクロックs t cのカウント値とビデオタイムスタンプV T S (V T S 1、V T S 2、V T S 3、... ..)との

10

20

30

40

50

間で、例えば時点 $T_v 1$ においてシステムタイムクロック $s t c$ のカウント値とビデオタイムスタンプ $V T S 1$ との差分値 $D 2 V$ が発生することはない。

【0121】

しかしながら、リアルタイムストリーミングエンコーダ 11 の PCR 回路 61 から供給されるクロックリファレンス $p c r$ はリアルタイム性が要求される UDP で送信されてくるものであり、高速性を重視するあまり再送制御されないので当該クロックリファレンス $p c r$ が第 2 のコンテンツ受信装置 4 のリアルタイムストリーミングデコーダ 12 へ到達しないか、あるいはエラーデータを含んで到達することもある。

【0122】

このような場合には、リアルタイムストリーミングエンコーダ 11 の PCR 回路 61 から所定周期で供給されてくるクロックリファレンス $p c r$ の値と、システムタイムクロック $s t c$ のクロック周波数との同期が PLL を介してずれることがあるが、このときも本発明におけるレンダラー 77 ではリップシンクを保障し得るようになされている。

10

【0123】

本発明では、システムタイムクロック $s t c$ とオーディオタイムスタンプ $A T S$ そしてビデオタイムスタンプ $V T S$ との間にずれが生じた場合、やはりリップシンクを取る方法として、オーディオ出力の連続性を優先させるようになされている。

【0124】

レンダラー 77 は、オーディオフィーム $A F 2$ の出力タイミング $T a 2$ でのシステムタイムクロック $s t c$ のカウント値とオーディオタイムスタンプ $A T S 2$ とを比較し、その差分値 $D 2 A$ を記憶する。一方、レンダラー 77 はビデオフレーム $V F 2$ の出力タイミング $T v 2$ でのシステムタイムクロック $s t c$ のカウント値とビデオタイムスタンプ $V T S 2$ とを比較し、その差分値 $D 2 V$ を記憶する。

20

【0125】

このとき、クロックリファレンス $p c r$ が第 2 のコンテンツ受信装置 4 のリアルタイムストリーミングデコーダ 12 へ確実に到達し、クロックリファレンス $p c r$ の値と当該リアルタイムストリーミングデコーダ 12 のシステムタイムクロック $s t c$ のクロック周波数とが PLL を介して完全に一致し、モニタ 13 を含んでデコーダ側がシステムタイムクロック $s t c$ に同期していれば差分値 $D 2 V$ 、 $D 2 A$ は「0」となる。

【0126】

この差分 $D 2 A$ が正値であればオーディオフィーム $A F 2$ は早いと判断され、負値であればオーディオフィーム $A F 2$ は遅れていると判断される。同様に、差分 $D 2 V$ が正値であればビデオフレーム $V F 2$ は早いと判断され、負値であればビデオフレーム $V F 2$ は遅れていると判断される。

30

【0127】

ここでレンダラー 77 は、オーディオフィーム $A F 2$ が早くても遅れていても、オーディオ出力の連続性を維持させることを優先させ、オーディオフィーム $A F 2$ に対するビデオフレーム $V F 2$ の出力を相対的に次のように制御する。

【0128】

例えば、時点 $T v 2$ のタイミングにおいて、 $D 2 V - D 2 A$ が閾値 $T H$ よりも大きい場合、差分値 $D 2 V$ が差分値 $D 2 A$ よりも大きければ音声に映像が追いついていない状態であるため、レンダラー 77 は GOP を構成している例えば B ピクチャに相当するビデオフレーム $V f 3$ (図示せず) をデコードすることなくスキップして次のビデオフレーム $V f 4$ を出力するようになされている。

40

【0129】

この場合、レンダラー 77 は出力ビデオバッファ 79 に格納されている「P」ピクチャについては、ビデオデコーダ 76 で次のピクチャをデコードする際の参照フレームとなるためスキップせず、次のピクチャを生成する際の参照フレームとならない非参照フレームである「B」ピクチャをスキップすることにより、画質劣化を未然に防ぎながらリップシンクさせるようになされている。

50

【0130】

これに対してD2V - D2Aが閾値THよりも大きく、差分値D2Aの方が差分値D2Vよりも大きければ映像に音声追いついていない状態であるため、レンダラー77は現在出力中のビデオフレームVf2を繰り返しリピート出力するようになされている。

【0131】

また、D2V - D2Aが閾値THよりも小さい場合は、音声に対する映像のギャップは許容範囲内であると判断され、レンダラー77は当該ビデオフレームVF2をそのままモニタ13へ出力する。

【0132】

ところでレンダラー77では、仮にスキップすべき「B」ピクチャが出力ビデオバッファ79に存在せず、「I」ピクチャや「P」ピクチャばかりであった場合には、当該「B」ピクチャをスキップすることはできないため、音声に映像を追い付かせることができなくなってしまう。

【0133】

そこでレンダラー77では、第1のコンテンツ受信装置3におけるストリーミングデコーダ9のレンダラー37と同様に、スキップすべき「B」ピクチャが存在しないときには、モニタ13のモニタ出力タイミングが例えば60 [Hz]であり、出力ビデオバッファ79から出力すべきビデオフレームVF2のピクチャリフレッシュタイミングが30 [Hz]であることを利用し、当該ピクチャリフレッシュタイミングを短縮するようになされている。

【0134】

具体的にはレンダラー77は、クロックリファレンスpcrと同期したシステムタイムクロックstcとビデオタイムスタンプVTSの差分値が16.666... [msec]を超えると、すなわち音声の出力タイミングに対してモニタ出力タイミングが1フレーム分以上遅れているときは、1フレーム分のビデオフレームをスキップする代わりにピクチャリフレッシュタイミングを30 [Hz]から60 [Hz]に変更して表示間隔を短縮するようになされている。

【0135】

つまりレンダラー77は、当該スキップによる画質劣化の影響を受ける「I」ピクチャや「P」ピクチャについてはピクチャリフレッシュ間隔を1/30秒から1/60秒に短縮することにより、「I」ピクチャや「P」ピクチャをスキップすることによる画質劣化を生じさせることなく映像を音声に追い付かせることができるようになされている。

【0136】

(7-2) ライブストリーミングにおけるリップシンク調整処理手順

上述のようにリアルタイムストリーミングデコーダ12のレンダラー77がライブストリーミング再生を行う際に、オーディオフレームAF2を基準にしてビデオフレームVF2の出力タイミングを調整することにより映像と音声とをリップシンクさせる出力タイミング調整方法についてまとめると、次の図14のフローチャートに示すように、リアルタイムストリーミングデコーダ12のレンダラー77は、ルーチンRT2の開始ステップから入って、次のステップSP21へ移る。

【0137】

ステップSP21において、第2のコンテンツ受信装置4におけるリアルタイムストリーミングデコーダ12のレンダラー77は、第1のコンテンツ受信装置3におけるリアルタイムストリーミングエンコーダ11のPCR回路61からクロックリファレンスpcrを受信し、次のステップSP22へ移る。

【0138】

ステップSP22においてレンダラー77は、減算回路81、フィルタ82、電圧制御型クリスタルオシレータ回路83及びSTC回路84を介して構成されるPLLによってクロックリファレンスpcrとシステムタイムクロックstcとを同期させることにより、これ以降、出力タイミングを調整する際の基準として当該クロックリファレンスpcrに同期したシステムタイムクロックstcを用い、次のステップSP23へ移る。

10

20

30

40

50

【 0 1 3 9 】

ステップ S P 2 3 においてレンダラー 7 7 は、時点 T v 1、T v 2、T v 3、... のタイミングにおけるシステムタイムクロック s t c のカウント値とビデオタイムスタンプ V T S との差分値 D 2 V を算出し、また時点 T a 1、T a 2、T a 3、... のタイミングにおけるシステムタイムクロック s t c のカウント値とオーディオタイムスタンプ A T S との差分値 D 2 A を算出し、次のステップ S P 2 4 へ移る。

【 0 1 4 0 】

ステップ S P 2 4 においてレンダラー 7 7 は、ステップ S P 2 3 で算出した差分値 D 2 V、D 2 A に基づいて算出した $D 2 V - D 2 A$ が閾値 T H (例えば 1 0 0 [msec]) よりも小さい場合、否定結果を得て次のステップ S P 2 5 へ移る。

10

【 0 1 4 1 】

ステップ S P 2 5 においてレンダラー 7 7 は、 $D 2 A - D 2 V$ が閾値 T H (例えば 1 0 0 [msec]) よりも大きい場合、肯定結果を得て映像が音声に対して進んでいると判断し、次のステップ S P 2 6 へ移る。

【 0 1 4 2 】

ステップ S P 2 6 においてレンダラー 7 7 は、映像の方が音声よりも進んでいるので、音声映像に追いつくように現在出力中のピクチャを構成するビデオフレーム V F 2 をリピートして出力した後、次のステップ S P 3 1 へ移って処理を終了する。

【 0 1 4 3 】

これに対してステップ S P 2 5 で $D 2 A - D 2 V$ が閾値 T H を越えていないのであれば、否定結果を得て、音声と映像との間にずれが生じているとは感じない程度であると判断し、次のステップ S P 2 7 へ移る。

20

【 0 1 4 4 】

ステップ S P 2 7 においてレンダラー 7 7 は、映像と音声との間でずれが生じているとは感じない程度の時間差しかないので、この場合はクロックリファレンス p c r と同期したシステムタイムクロック s t c を基に、ビデオフレーム V F 2 をビデオタイムスタンプ V T S に従ってそのままモニタ 1 3 へ出力し、次のステップ S P 3 1 へ移って処理を終了する。

【 0 1 4 5 】

なおレンダラー 7 7 は、音声に関しては音の連続性を維持させるため、上記のいずれの場合においても、クロックリファレンス p c r と同期したシステムタイムクロック s t c を基に、オーディオタイムスタンプ A T S に従ってそのままモニタ 1 3 へ出力するようになされている。

30

【 0 1 4 6 】

これに対してステップ S P 2 4 で肯定結果が得られると、このことは $D 2 V - D 2 A$ が閾値 T H (例えば 1 0 0 [msec]) よりも大きいこと、すなわち音声に対して映像が遅れていることを表しており、このときレンダラー 7 7 は次のステップ S P 2 8 へ移る。

【 0 1 4 7 】

ステップ S P 2 8 においてレンダラー 7 7 は、出力ビデオバッファ 7 9 に「B」ピクチャが存在するか否かを判定し、肯定結果が得られると次のステップ S P 2 9 へ移り、否定結果が得られると次のステップ S P 3 0 へ移る。

40

【 0 1 4 8 】

ステップ S P 2 9 においてレンダラー 7 7 は、ビデオがオーディオに対して遅れていると判断し、かつ「B」ピクチャが出力ビデオバッファ 7 9 に存在することを確認したので、Bピクチャ(ビデオフレーム V f 3)をデコードせずにスキップして出力することにより、音声に対する映像の遅れを取り戻してリップシンクさせることができ、次のステップ S P 3 1 へ移って処理を終了する。

【 0 1 4 9 】

一方、ステップ S P 3 0 においてレンダラー 7 7 は、モニタ 1 3 のモニタ出力タイミングが 6 0 [Hz]であるのに対し、ビデオフレーム V F 2 のピクチャリフレッシュタイミング

50

が 30 [Hz]であることを利用し、当該ピクチャリフレッシュタイミングをモニタ 13 のモニタ出力タイミングに合わせて短縮することにより、ピクチャをスキップすることによる画質劣化を生じさせることなく音声に映像を追い付かせ、次のステップ S P 31 へ移って処理を終了する。

【0150】

このように第 2 のコンテンツ受信装置 4 におけるリアルタイムストリーミングデコーダ 12 のレンダラ 77 は、第 1 のコンテンツ受信装置 3 におけるリアルタイムストリーミングエンコーダ 11 のクロックリファレンス p c r と当該リアルタイムストリーミングデコーダ 12 のシステムタイムクロック s t c とを同期させることによりライブストリーミング再生を実現すると共に、そのためのクロックリファレンス p c r が U D P でリアルタイム性を重要視するために再送制御されずに到達しないことがあった場合でも、システムタイムクロック s t c に対するオーディオタイムスタンプ A T S、ビデオタイムスタンプ V T S のずれに応じてリップシンク調整処理を実行することにより、ライブストリーミング再生を行いながらも確実にリップシンクし得るようになされている。

10

【0151】**(8) 動作及び効果**

以上の構成において、第 1 のコンテンツ受信装置 3 のストリーミングデコーダ 9 は、コンテンツの種類がオーディオ及びビデオからなるものである場合、プリセット用ビデオタイムスタンプ V T S p でシステムタイムクロック s t c の値をプリセットした後に必ずプリセット用オーディオタイムスタンプ A T S p でプリセットし直すことにより、最終的には必ずシステムタイムクロック s t c の値とプリセット用オーディオタイムスタンプ A T S p (A T S p 1、A T S p 2、A T S p 3、...) とを一致させる。

20

【0152】

ストリーミングデコーダ 9 のレンダラ 37 は、プリセット用オーディオタイムスタンプ A T S p でプリセットしたシステムタイムクロック s t c のカウント値と、ビデオフレーム V F 1 (V f 1、V f 2、V f 3、...) に付されたビデオタイムスタンプ V T S (V T S 1、V T S 2、V T S 3、...) との差分値 D 1 を算出することにより、当該ビデオタイムスタンプ V T S を付したエンコーダ側のクロック周波数とデコーダ側システムタイムクロック s t c のクロック周波数とのずれによって生じる時間差を認識することができる。

30

【0153】

ストリーミングデコーダ 9 のレンダラ 37 は、その差分値 D 1 に応じてビデオフレーム V F 1 の現ピクチャをリピートして出力したり、又は非参照フレームの B ピクチャをデコードせずにスキップして出力したり、或いはピクチャリフレッシュタイミングを短縮して出力することにより、モニタ 10 へ出力する音声を途切れさせることなく当該音声の連続性を保ったまま、その音声に対する映像の出力タイミングを調整することができる。

【0154】

もちろんレンダラ 37 は、差分値 D 1 が閾値 T H 以下であって、ユーザがリップシンクのずれを認識し得ない程度である場合には、リピート出力やスキップ再生処理或いはピクチャリフレッシュ間隔を短縮することをせずにビデオタイムスタンプ V T S (V T S 1、V T S 2、V T S 3、...) の通りにモニタ 10 へ出力することもできるので、この場合にも当然映像の連続性を保つことができる。

40

【0155】

さらに第 2 のコンテンツ受信装置 4 におけるリアルタイムストリーミングデコーダ 12 のレンダラ 77 は、第 1 のコンテンツ受信装置 3 におけるリアルタイムストリーミングエンコーダ 11 の P C R 回路 61 から供給されるクロックリファレンス p c r とデコーダ側のシステムタイムクロック s t c とを同期させた上で、オーディオタイムスタンプ A T S 及びビデオタイムスタンプ V T S に従ってオーディオフレーム A F 2 及びビデオフレーム V F 2 をモニタ 13 へ出力することができるので、リアルタイム性を保持したままライブストリーミング再生を実現することができる。

50

【0156】

その上、第2のコンテンツ受信装置4におけるリアルタイムストリーミングデコーダ12のレンダラー77は、第1のコンテンツ受信装置3におけるリアルタイムストリーミングエンコーダ11のPCR回路61から供給されるクロックリファレンス p_{cr} がUDPで再送制御されずに到達しないために、当該クロックリファレンス p_{cr} とシステムタイムクロック s_{tc} との同期が外れたとしても、システムタイムクロック s_{tc} とビデオタイムスタンプ VTS との差分値 $D2V$ 、システムタイムクロック s_{tc} とオーディオタイムスタンプ ATS との差分値 $D2A$ を算出し、当該差分値 $D2V$ と $D2A$ とのギャップに応じてビデオフレーム $VF2$ の出力タイミングを調整することにより、モニタ13へ出力する音声を途切れさせることなく連続性を保ったまま、その音声に対する映像の出力タイミングを調整することができる。

【0157】

また第1のコンテンツ受信装置3におけるストリーミングデコーダ9のレンダラー37は、プリセット用ビデオタイムスタンプ VTS_p 、プリセット用オーディオタイムスタンプ ATS_p の順番で決められたプリセットシーケンスに従い、当該プリセット用ビデオタイムスタンプ VTS_p 、プリセット用オーディオタイムスタンプ ATS_p を用いてシステムタイムクロック s_{tc} をプリセットするようにしたことにより、コンテンツの種類がオーディオだけからなるものであるときにはプリセット用オーディオタイムスタンプ ATS_p でシステムタイムクロック s_{tc} をプリセットすることができ、またコンテンツの種類がビデオのみでなるものであるときには、プリセット用ビデオタイムスタンプ VTS_p でシステムタイムクロック s_{tc} をプリセットすることができるので、コンテンツの種類がオーディオ及びビデオからなるもの、オーディオのみからなるもの又はビデオのみからなるものである場合にも対応することができる。

【0158】

すなわちストリーミングデコーダ9のレンダラー37は、コンテンツが必ずしもオーディオ及びビデオからなるものでなく、当該コンテンツがビデオのみからなるものであってプリセット用オーディオタイムスタンプ ATS_p が存在しないときや、コンテンツがオーディオのみからなるものであってプリセット用ビデオタイムスタンプ VTS_p が存在しないときでも、ビデオフレーム $VF1$ やオーディオフレーム $AF1$ の出力に対応することができるので、コンテンツの種類に応じた最適なタイミングでモニタ10へ出力することができる。

【0159】

さらにストリーミングデコーダ9のレンダラー37では、プリセット後のシステムタイムクロック s_{tc} のカウント値とビデオタイムスタンプ $VTS2$ との差分値 $D1$ が所定の閾値 TH よりも大きく、かつ映像が音声よりも遅れている場合で、出力ビデオバッファ39にBピクチャが存在するときには画質劣化の影響のない当該Bピクチャをデコードすることなくスキップし、出力ビデオバッファ39にBピクチャが存在しないときにはモニタ10のモニタ出力タイミングに合わせてビデオフレーム $VF1$ のピクチャリフレッシュタイミングを短縮することによりピクチャスキップによる画質劣化を生じさせずに映像を音声に追い付かせることができる。

【0160】

以上の構成によれば、第1のコンテンツ受信装置3におけるストリーミングデコーダ9のレンダラー37及び第2のコンテンツ受信装置4におけるリアルタイムストリーミングデコーダ12のレンダラー77は、オーディオフレーム $AF1$ 、 $AF2$ の出力タイミングを基準としてビデオフレーム $VF1$ 、 $VF2$ の出力タイミングを調整することができるので、音声の連続性を保ったまま視聴者であるユーザに違和感を感じさせることなくリップシンクさせることができる。

【0161】

(9) 他の実施の形態

なお上述の実施の形態においては、オーディオフレーム $AF1$ 、 $AF2$ を基準とした差

10

20

30

40

50

分値 D 1 又は D 2 V、D 2 A に応じてリップシンクを調整することによりエンコーダ側のクロック周波数とデコーダ側のクロック周波数とのずれを吸収するようにした場合について述べたが、本発明はこれに限らず、クロックジッタ、ネットワークジッタ等によって生じるエンコーダ側のクロック周波数とデコーダ側のクロック周波数との微妙なずれを吸収するようにしても良い。

【 0 1 6 2 】

また上述の実施の形態においては、コンテンツ提供装置 2 と第 1 のコンテンツ受信装置 3 との間でインターネット 5 を介して接続し、プリエンコードストリーミングを実現するようにした場合について述べたが、本発明はこれに限らず、コンテンツ提供装置 2 と第 2 のコンテンツ受信装置 4 との間でインターネット 5 を介して接続し、プリエンコードストリーミングを実現するようしたり、コンテンツ提供装置 2 から第 1 のコンテンツ受信装置 3 を介して第 2 のコンテンツ受信装置 4 へコンテンツを提供することによりプリエンコードストリーミングを実現するようにしても良い。

10

【 0 1 6 3 】

さらに上述の実施の形態においては、第 1 のコンテンツ受信装置 3 と第 2 のコンテンツ受信装置 4 との間でライブストリーミングを行うようにした場合について述べたが、本発明はこれに限らず、コンテンツ提供装置 2 と第 1 のコンテンツ受信装置 3 との間や、コンテンツ提供装置 2 と第 2 のコンテンツ受信装置 4 との間でライブストリーミングを行うようにしても良い。

【 0 1 6 4 】

この場合、コンテンツ提供装置 2 のストリーミングサーバ 8 からクロックリファレンス p c r を第 1 のコンテンツ受信装置 3 のストリーミングデコーダ 9 へ送信し、当該ストリーミングデコーダ 9 でクロックリファレンス p c r とシステムタイムクロック s t c とを同期させることにより、ライブストリーミングを実現することができる。

20

【 0 1 6 5 】

さらに上述の実施の形態においては、サブトラクタ回路 4 4 及び 4 5 によって最初のビデオタイムスタンプ V T S 1 及びオーディオタイムスタンプ A T S 1 の値を所定時間分だけ引き戻すようにした場合について述べたが、本発明はこれに限らず、バッファ 4 2 及び 4 3 を介して最初のビデオタイムスタンプ V T S 1 及びオーディオタイムスタンプ A T S 1 がコンパレータ回路 4 6 に到達した時点で、バッファ 4 2 及び 4 3 による遅延等により S T C 回路 4 1 からコンパレータ回路 4 6 に供給されるプリセット後のシステムタイムクロック s t c の値が当該ビデオタイムスタンプ V T S 1 及びオーディオタイムスタンプ A T S 1 を経過していることがなければ、必ずしもサブトラクタ回路 4 4 及び 4 5 によって最初のビデオタイムスタンプ V T S 1 及びオーディオタイムスタンプ A T S 1 の値を所定時間分だけ引き戻さなくても良い。

30

【 0 1 6 6 】

さらに上述の実施の形態においては、コンテンツの種類を判別する前に当該コンテンツの種類に拘らずプリセット用ビデオタイムスタンプ V T S p、プリセット用オーディオタイムスタンプ A T S p の順番で決められたプリセットシーケンスに従ってシステムタイムクロック s t c をプリセットするようにした場合について述べたが、本発明はこれに限らず、最初にコンテンツの種類を判別し、当該コンテンツがオーディオ及びビデオからなるものであるときには、プリセット用ビデオタイムスタンプ V T S p、プリセット用オーディオタイムスタンプ A T S p でシステムタイムクロック s t c をプリセットし、当該コンテンツがビデオだけからなるものであるときには、プリセット用ビデオタイムスタンプ V T S p でシステムタイムクロック s t c をプリセットし、当該コンテンツがオーディオだけからなるものであるときには、プリセット用オーディオタイムスタンプ A T S p でシステムタイムクロック s t c をプリセットするようにしても良い。

40

【 0 1 6 7 】

さらに上述の実施の形態においては、出力ビデオバッファ 3 9、7 9 に B ピクチャが存在しない場合には、モニタ 1 0 及び 1 3 のモニタ出力レートに合わせてビデオフレーム V

50

F 1 及び V F 2 のピクチャリフレッシュレートを 3 0 [Hz] から 6 0 [Hz] に短縮するようにした場合について述べたが、本発明はこれに限らず、B ピクチャの有無に拘らずビデオフレーム V F 1 及び V F 2 のピクチャリフレッシュレートを 3 0 [Hz] から 6 0 [Hz] に短縮するようにしても良い。この場合でも、レンダラ 3 7 及び 7 7 は、音声に対して映像の遅れを取り戻してリップシンクさせることができる。

【 0 1 6 8 】

さらに上述の実施の形態においては、B ピクチャをスキップして出力するようにした場合について述べたが、本発明はこれに限らず、I ピクチャの直前に位置する P ピクチャをスキップして出力するようにしても良い。

【 0 1 6 9 】

これは、I ピクチャの直前に位置する P ピクチャであれば、次の I ピクチャを生成する際に当該 P ピクチャが参照されることはなく、スキップしたとしても次の I ピクチャを生成する際に支障を来すことなく、画質劣化が生じることもないからである。

【 0 1 7 0 】

さらに上述の実施の形態においては、ビデオフレーム V f 3 をデコードせずにスキップしてモニタ 1 0 へ出力するようにした場合について述べたが、本発明はこれに限らず、ビデオフレーム V f 3 をデコードした後に出力ビデオバッファ 3 9 から出力する段階でデコード後のビデオフレーム V f 3 をスキップして出力するようにしても良い。

【 0 1 7 1 】

さらに上述の実施の形態においては、オーディオフレーム A F 1、A F 2 についてはリップシンクの調整を行う際の基準として用いているために、全てのオーディオフレームについて欠けることなくモニタ 1 0、1 3 へ出力するようにした場合について述べたが、本発明はこれに限らず、例えば無音部分に相当するオーディオフレームがあった場合には、そのオーディオフレームをスキップして出力するようにしても良い。

【 0 1 7 2 】

さらに上述の実施の形態においては、本発明のコンテンツ受信装置を、復号手段としてのオーディオデコーダ 3 5、7 4、ビデオデコーダ 3 6、7 6 と、記憶手段としての入力オーディオバッファ 3 3、7 3、出力オーディオバッファ 3 8、7 8、入力ビデオバッファ 3 4、7 5、出力ビデオバッファ 3 9、7 9 と、算出手段及びタイミング調整手段としてのレンダラ 3 7、7 7 とによって構成するようにした場合について述べたが、本発明はこれに限らず、その他種々の回路構成でコンテンツ受信装置を形成するようにしても良い。

【 産業上の利用可能性 】

【 0 1 7 3 】

本発明のコンテンツ受信装置、ビデオオーディオ出力タイミング制御方法及びコンテンツ提供システムは、例えばサーバから音声付の動画コンテンツをダウンロードして表示する用途に適用することができる。

【 図面の簡単な説明 】

【 0 1 7 4 】

【 図 1 】 ストリーミングシステムの全容を表すコンテンツ提供システムの全体構成を示す略線的ブロック図である。

【 図 2 】 コンテンツ提供装置の回路構成を示す略線的ブロック図である。

【 図 3 】 オーディオパケット及びビデオパケット内のタイムスタンプ (T C P プロトコル) の構造を示す略線図である。

【 図 4 】 第 1 のコンテンツ受信装置におけるストリーミングデコーダのモジュール構成を示す略線的ブロック図である。

【 図 5 】 タイミングコントロール回路の構成を示す略線的ブロック図である。

【 図 6 】 プリセット後の S T C と比較されるタイムスタンプを示す略線図である。

【 図 7 】 プリエンコードストリーミングにおけるビデオフレーム及びオーディオフレームの出力タイミングを説明する際に供する略線図である。

10

20

30

40

50

【図 8】I ピクチャ、P ピクチャの場合のビデオフレーム出力制御処理の説明に供する略線図である。

【図 9】プリエンコードストリーミングにおけるリップシンク調整処理手順を示すフローチャートである。

【図 10】第 1 のコンテンツ受信装置におけるリアルタイムストリーミングエンコーダの回路構成を示す略線的ブロック図である。

【図 11】コントロールパケット内の PCR (UDP プロトコル) の構造を示す略線図である。

【図 12】第 2 のコンテンツ受信装置におけるリアルタイムストリーミングデコーダの回路構成を示す略線的ブロック図である。

10

【図 13】ライブストリーミングにおけるビデオフレーム及びオーディオフレームの出力タイミングを説明する際に供する略線図である。

【図 14】ライブストリーミングにおけるリップシンク調整処理手順を示す略線的フローチャートである。

【符号の説明】

【0175】

1 …… コンテンツ提供システム、2 …… コンテンツ提供装置、3 …… 第 1 のコンテンツ受信装置、4 …… 第 2 のコンテンツ受信装置、5 …… インターネット、7 …… エンコーダ、8 …… ストリーミングサーバ、9 …… ストリーミングデコーダ、10、13 …… モニタ、11 …… リアルタイムストリーミングエンコーダ、12 …… リアルタイムストリーミングデコーダ、14 …… Webサーバ、15 …… Webブラウザ、21、51 …… ビデオ入力部、22、52 …… ビデオエンコーダ、23 …… ビデオES蓄積部、24、53 …… オーディオ入力部、25、54 …… オーディオエンコーダ、26 …… オーディオES蓄積部、28、57 …… ビデオフレームカウンタ、29、58 …… オーディオフレームカウンタ、27、56 …… パケット生成部、30、59 …… パケットデータ蓄積部、31、71 …… 入力パケット蓄積部、32、72 …… パケット分割部、33、73 …… 入力オーディオバッファ、34、75 …… 入力ビデオバッファ、35、74 …… オーディオデコーダ、36、76 …… ビデオデコーダ、37、77 …… レンダラー、38、78 …… 出力オーディオバッファ、39、79 …… 出力ビデオバッファ、40 …… クリスタルオシレータ回路、81 …… 減算回路、82 …… フィルタ、83 …… 電圧制御型クリスタルオシレータ、41、60、84 …… STC回路、61 …… PCR回路。

20

30

【 図 1 】

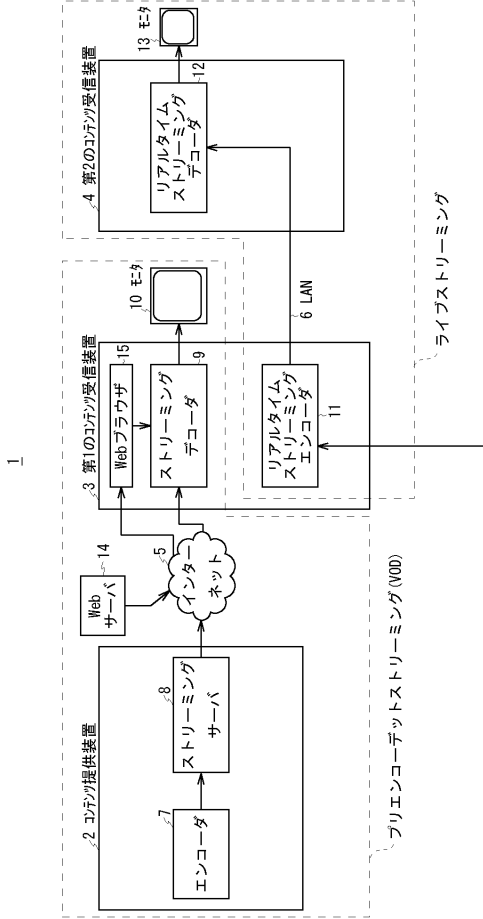


図 1 ストリーミングシステムの全体を表したコンテンツ提供システム

【 図 2 】

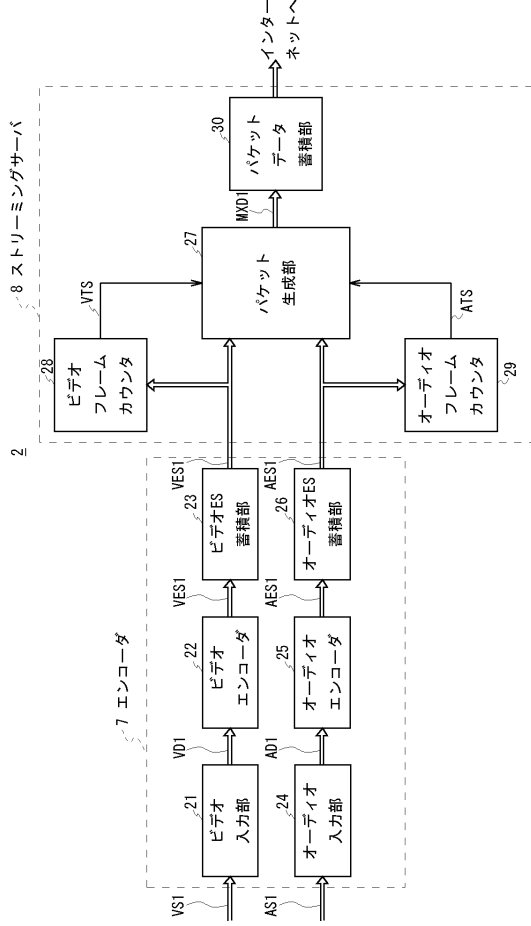


図 2 コンテンツ提供装置の回路構成

【 図 3 】

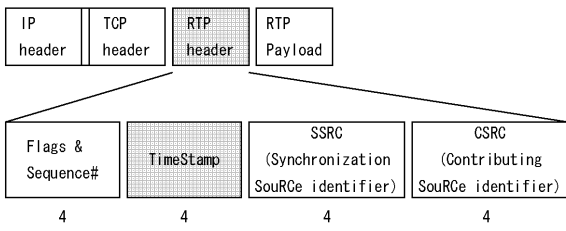


図 3 オーディオパケット及びビデオパケット内のタイムスタンプ (TCPプロトコル)

【 図 4 】

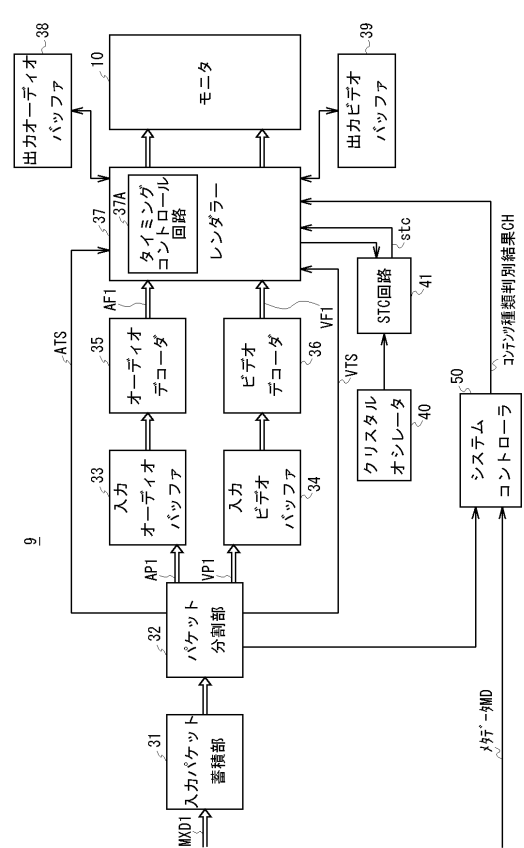


図 4 第 1 のコンテンツ受信装置におけるストリーミングデコーダのモジュール構成

【 図 5 】

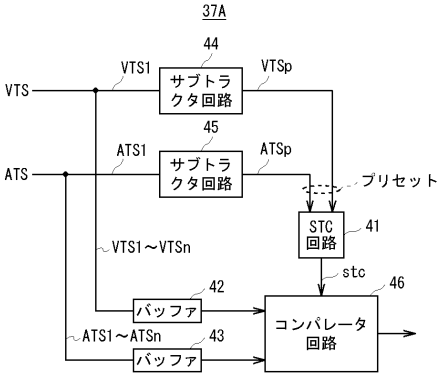


図5 タイミングコントロール回路の構成

【 図 6 】

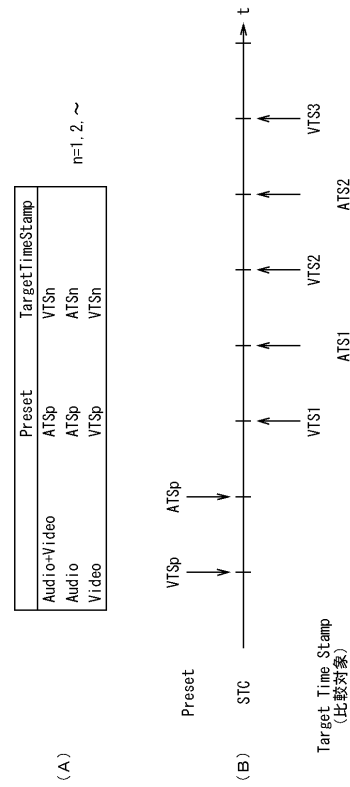


図6 プリセット後のSTCと比較されるタイムスタンプ

【 図 7 】

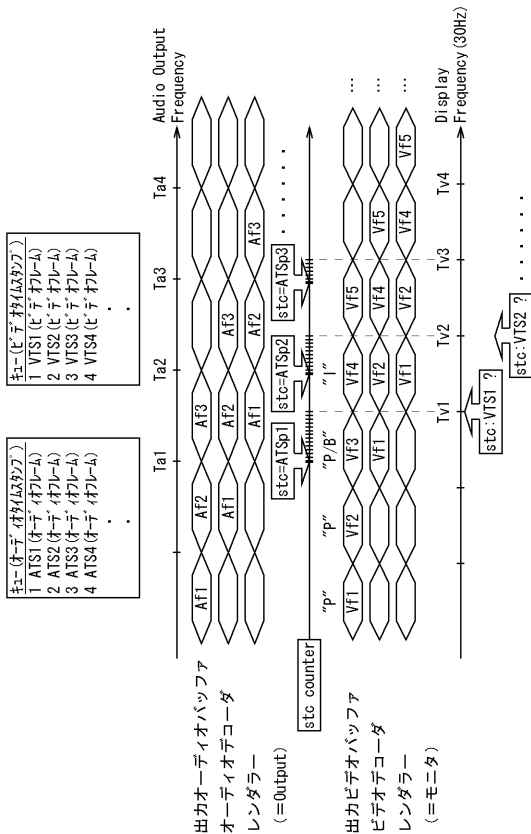


図7 プリエンコーデッドストリーミングにおけるビデオフレーム及びオーディオフレームの出カタイミミング

【 図 8 】

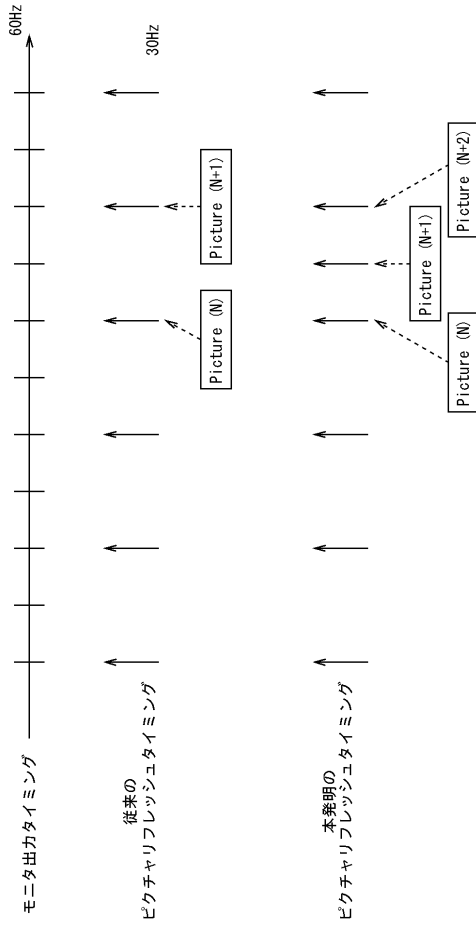


図8 Iピクチャ、Pピクチャの場合のビデオフレーム出力制御処理

【 図 9 】

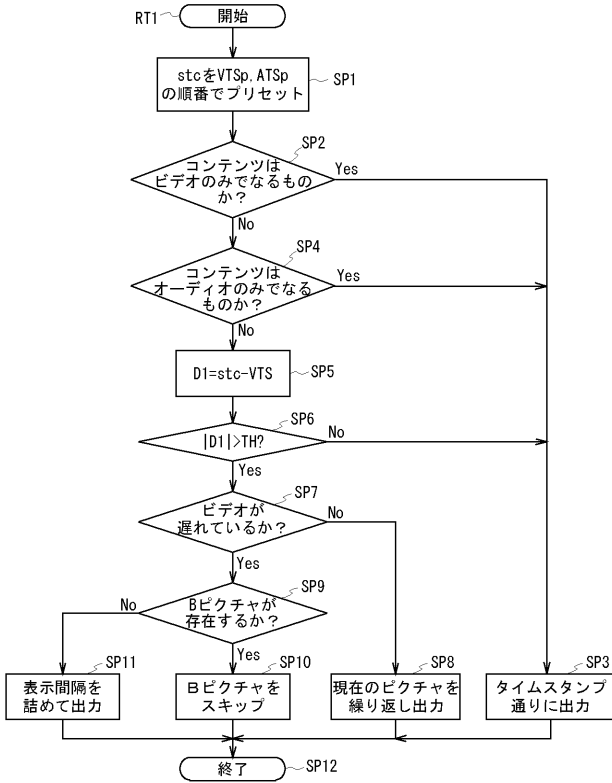


図9 プリエンコードドストリーミングにおけるリップシンク調整処理手順

【 図 10 】

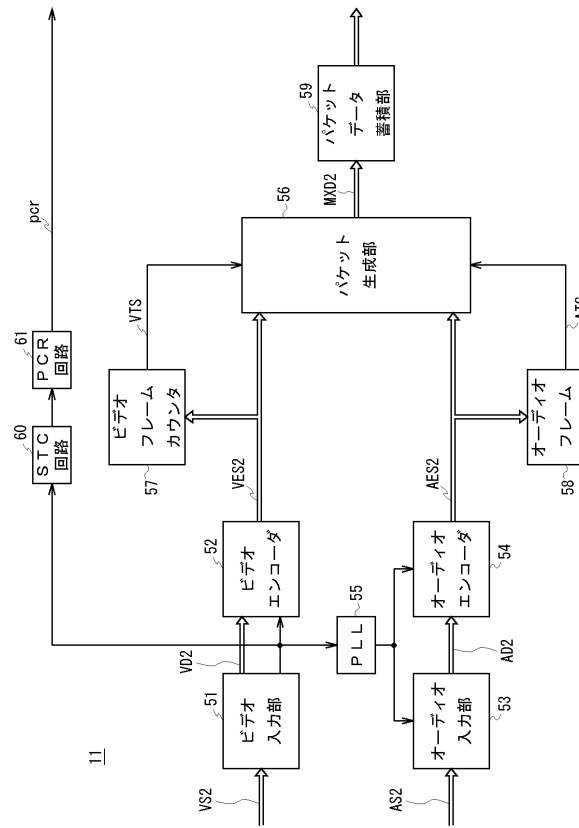


図10 第1のコンテンツ受信装置におけるリアルタイムストリーミングエンジン回路構成

【 図 11 】

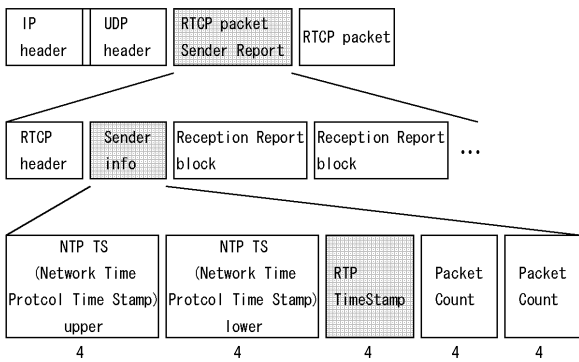


図11 コントロールパケット内のPCR (UDFプロトコル)

【 図 12 】

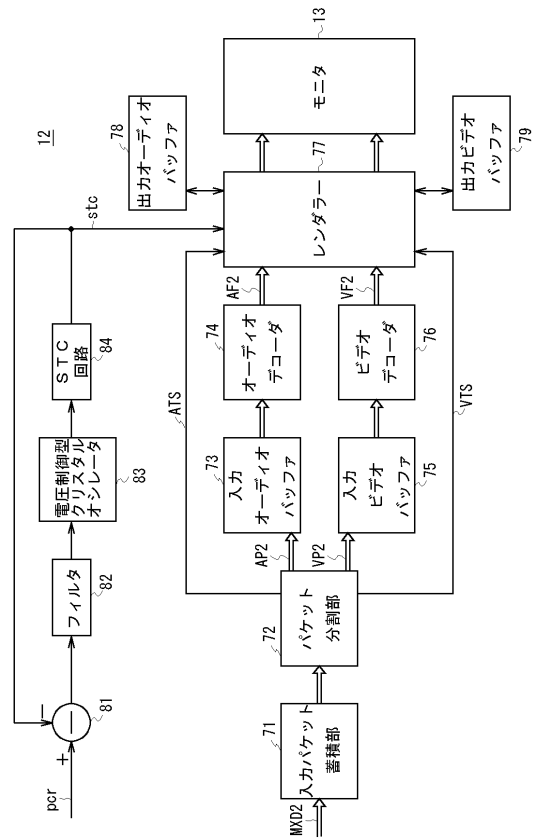


図12 第2のコンテンツ受信装置におけるリアルタイムストリーミングデコーダ回路構成

フロントページの続き

Fターム(参考) 5C053 FA28 GB06 GB37 HC05 JA03 JA22 LA14
5C059 RC03 RC04 RC32 SS30 UA02 UA05

【要約の続き】

【選択図】 図7